

CC diagnostic markers and pharmaceutical agents for various diseases of the
CC central and peripheral nervous system, as well as cardiac, urologic and
CC gastrointestinal disorders.

SQ Sequence 1467 BP; 258 A; 459 C; 432 G; 306 T; 12 other;

Query Match 72.0%; Score 1310.2; DB 21; Length 1467;
Best Local Similarity 96.9%; Pred. No. 2.2e-289;
Matches 1423; Conservative 7; Mismatches 27; Indels 12; Gaps 9;

Matches 1423: Conservative 7; Mismatches 27; Indels 12; Gaps 9;									
Qy	195	gtactacaacctgtgtgaccgctctgggcgctggggcattccctggaggccgtggctgg	254						
Db	1	gtactacaacctgtgtgaccgctctgggcgctggggcattccctggaggccgtggctgg	60						
Qy	255	ggcggggaattgtaccacagtttgtctcaacatcatctgtgtggccagctccccccttgt	314						
Db	61	ggcgggcaattgtaccacagtttgtctcaacatcatctgtgtggccagctccccccttgt	120						
Qy	315	gcagagaccacgaagagccctgtcgggagaccaggtattctctctcctcctcctcctc	374						
Db	121	gcaggacaccgaagagccctgtcgggagaccaggtattctctctcctcctcctcctc	180						
Qy	375	ggggctc---tctcgcctcgtgttgcctgtgtgtggtgaagcccgac-tctccaccctgt	430						
Db	181	tggggcctcttctgcctcgtgttgcctgtkktggtgaagcccgaccttctccaccctgt	240						
Qy	431	cctctcggcgctctcctcttctggggttctgttcggccatctgtcttcttctgtctggc	490						
Db	241	cctctcggcgctctcctcttctggggttctgttcggccatctgtcttcttctgtctggc	300						
Qy	491	acgtcttttgcctcaacttctcggccgggaagaccacggcccgggctgtggtgatct	550						
Db	301	acgtcttttgcctcaacttctcggccgggaagaccacggcccgggctgtggtgatct	360						
Qy	551	tcactgtggctctgctctgacctgttagaggttcatacatcaatacacagatggctgatca	610						
Db	361	tcactgtggctctgctctgacctgttagaggttcatacatcaatacacagatggctgatca	420						
Qy	611	tcaccttggttcggggcagtgcgaggggcgccctcagggcaacagcagcgacggctggg	670						
Db	421	tcaccttggttcggggcagtgcgaggggcgccctcagggcaacagcagcgacggctggg	480						
Qy	671	ccgtggcctccccctgtgccatcgccaacatgggactttgtcatggcactcatctacgtca	730						
Db	481	ccgtggcctccccctgtgccatcgccaacatgggactttgtcatggcactcatctacgtca	540						
Qy	731	tgtcgtcgtcgtcgtcgtcctctcctcggggcgctggcccgccctgtgtgcgcgtacaa	790						
Db	541	tgtcgtcgtcgtcgtcgtcctcctcggggcgctggcccgccctgtgtgcgcgtacaa	600						
Qy	791	gctggcgttaagcataggggctcttctcctccacacagccacctcggttgccatatggg	850						
Db	601	gctggcgttaagcataggggctcttctcctccacacagccacctcggttgccatatggg	660						
Qy	851	tgggtgtggaatcgatgatacttaccggcaacagcagcacaacagtcctccaccctgggatg	910						
Db	661	tgggtgtggaatcgatgatacttaccggcaacagcagcacaacagtcctccaccctgggatg	720						
Qy	911	acccacagctggccatcgccctcgcgcgaatgctgtgggcttgcgtctctcttctacgtca	970						
Db	721	acccacagctggccatcgccctcgcgcgaatgctgtgggcttgcgtctctcttctacgtca	780						
Qy	971	tccccgaggctcccgagtgaccaagtccagccagacgaacagctaccagggggacatgt	1030						
Db	781	tccccgaggctcccgagtgaccaagtccagccagacgaacagctaccagggggacatgt	840						
Qy	1031	acccacccccgggg-cgtggggctatgagacacatcctgaagagcagaagggctcagagcatg	1089						
Db	841	acccacccccggggcgttgggctatgagacacatcctgaagagcagaaggggtcagagcatg	900						
Qy	1090	ttcgttgagacaagggccttttccatggatgagccggttgcagctaaagagccgggtgca	1149						

Db	901	ttctgaggaaacaaagcccttttccatggtgatgagccggttgacgataagaggccgggtgtca	960
Qy	1150	ccatacagcgggtacaatgggcagctgctgaccagtggtaccagcccaactgagatggcc	1209
Db	961	ccatacagcgggtacaaaggcagctgctgaccagtggtaccagcccaactgagatggcc	1020
Qy	1210	ctgatgcacaagtccgtccgaaggagcttacgacatcatctcccacgggcccacgcgc	1269
Db	1021	ctgatgcacaagtccgtccgaaggagctttacgacatcatctcccacgggcccacgcgc	1080
Qy	1270	aacagccagggtgatgggcagtgccaactcgaccctcgggctggaagacatgtactcgcc	1329
Db	1081	aacagccagggtgatgggcagtg-caactcgaccctgc-ggctgaagacatgtactcgcc	1138
Qy	1330	cagagccaccagcgccacacccgcgaagacggaagacggaagaaactctcaggtcttt-aaaa	1388
Db	1139	cagagccaccaggggaccacacgcgagaaagacggaagaaactctcaggtcttttwagaa	1198
Qy	1389	cccttcacgtgtgggactgagtcagcgggtggcga-ggaagagcgggtcggattt-ggggagg	1446
Db	1199	cccttcacgtgtgggactgagtcagcgggtggcrakgagagcgggtcggatttgggagg	1258
Qy	1447	gcctgaggacctggcccggggcaa--gggactctcaggctctctctcccctcggcagg	1504
Db	1259	gscctgaggacctggcccggggcaaaggatyyttccaggyttctctcccctcggcagg	1318
Qy	1505	ccagcaacatgtcccagatgtggaaggccctccctctctgcagtggttgggtgggt	1564
Db	1319	ccagcaacatgtgcccagatgtggaaggccctccctctctgcagtggttgggtgggt	1378
Qy	1565	gtcatgggtgtcccccaccactctctcagtggttggagtcgagagcgaacccacgcct	1624
Db	1379	gtcaatggtgtcccccaccactctctcagtggttggagtcgagagcgaacccacgcct	1438
Qy	1625	cctgcaggatcaactcgcggtcacact	1653
Db	1439	cctgcaggatcaactcgcggtcacact	1467
RESULT 15			
AAx98029			
ID	AAx98029 standard; DNA; 1034 BP.		
AC	AAx98029;		
XX			
DT	17-SEP-1999 (first entry)		
XX			
DE	Human secreted protein gene 79.		
XX			
KW	Human; secreted protein; cancer; tumour; developmental abnormality;		
KW	fetal deficiency; blood disorder; immune system disorder; inflammation;		
KW	autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;		
KW	schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;		
KW	athropclerosis; diabetes; cardiovascular disorder; kidney disorder;		
KW	digestive disorder; endocrine disorder; infection; AIDS; ss.		
OS	Homo sapiens.		
XX			
PN	W09931117-A1.		
XX			
PD	24-JUN-1999.		
XX			
PF	17-DEC-1998;	98WO-US27059.	
XX			
PR	19-DEC-1997;	97US-0068369.	
PR	18-DEC-1997;	97US-0068006.	
PR	18-DEC-1997;	97US-0068007.	
PR	18-DEC-1997;	97US-0068008.	
PR	18-DEC-1997;	97US-0068003.	
PR	18-DEC-1997;	97US-0068004.	
PR	18-DEC-1997;	97US-0068005.	
PR	18-DEC-1997;	97US-0068006.	
PR	18-DEC-1997;	97US-0070923.	

CC appetite, and can act as antigens in vaccines to raise an immune response
CC to the protein or another material cross-reactive with the protein.

Sequence 1323 BP; 232 A; 431 C; 384 G; 276 T; 0 other;
xx
SQ

Query Match 72.7%; Score 1323; DB 22; Length 1323;
Best Local Similarity 100.0%; Pred. No. 2.6e-292;

Qy	82	a t g g c a t c c a a a g c c t t g t g a t g t c g c t y g g a c t g c t c t t c c t t g t t c c a g g g	141
Db	1	a t g g c a t c c a a a g c c t t g t g a t g t c g c t y g g a c t g c t c t c t t c c t g t t c c a g g g	60
Qy	142	g c c t g g c c c a a g g c c a t c c c a c c g g c t g c a g c a a g c c t c a a c c c c t g t a c t a c	201
Db	61	g c c t g g c c c a a g g c c a t g t c c c a c c g g c t g c a g c a a g g c c t c a a c c c c t g t a c t a c	120
Qy	202	a a c c t g t g a c c g c t c t g g g c t g t g g g c a t g c t c c t g g a g g c c g t g g c t g g g g c g g c	261
Db	121	a a c c t g t g a c c g c t c t g g g c g t g g g c a t g c t c c t g g a g g c c g t g g c t g g g g c g g c	180
Qy	262	a t t g t c a c a a c g t t t g t c a c c a t c a t c c t g t g g c c a g c t c c c c t t t g t g a g a c	321
Db	181	a t t g t c a c a c g t t t g t c a c c a t c a t c c t g t g g c c a g c c c c c t t g t g a g a c	240
Qy	322	a c a a g a a a c g a g c c t g c t g g g g a c c a g g t a t t c t c c t c t g g a g a c c t g g g c t c	381
Db	241	a c a a g a a a c g a g c c t g c t g g g g a c c a g g t a t t c t c c t c t t c t g g a c c c t g g g c t c	300
Qy	382	t t c t g c c t g t t t g c c t g t g t g a a g c c g a c t t c t c a c c t g t g c e t c t g g c g c	441
Db	301	t t c t g c c t g t t t g c c t g t g t g a a g c c g a c t c t c a c c t g t g c t c t g g c g c	360
Qy	442	t t c c t t t t g g g t c t c g t t o g c a t c t g c t c t c t c t c t g g c g g c t a c g t t t g c c	501
Db	361	t t c c t t t t g g g t c t c t g t t g c a t c t g c t c t c t t g t g g g g c t a c g t t t g c c	420
Qy	502	c t c a a c t c t g c c g a a g a a c a c a g g g c c c g g g c t g g g t g a t c t c a c t g t g c t	561
Db	421	c t c a a c t c t c c g c c g a a g a a c a c a g g g c c c g g g c t g g g t g a t c t c a c t g t g g c t	480
Qy	562	c t g c t g t a c c c t g t a g a g t g c a t c a t c a a c a g a g t g g c t g a t c a c a c c c t g g t	621
Db	481	c t g c t g t a c c c t g t a g a g t g c a t c a t c a a t a c a g a g t g g c t g a t c a c c c t g g t	540
Qy	622	c g g g g c a g t g g c a g g g c g c c t c a g g g c a a c a g c a g c g t g g g c g t g g c c t c c	681
Db	541	c g g g g c a g t g g c a g g g c g c c t c a g g g c a a c a g c a g c g t g g g c g t g g c c t c c	600
Qy	682	c c c t g t g c a t c c c a a c a t g a c t t t g c a t g g c a c t c a t c a c t a c g t a c t g c t g c t g	741
Db	601	c c c t g t g c a t c c c a a c a t g a c t t t g c a t g g c a c t a c t a c g t a c t g c t g c t g c t g	660
Qy	742	c t g g g t g c c t c c t g g g g c c t g g c c g c c c t g t g t g c c t a c a a g c g t g g c g t a a g	801
Db	661	c t g g g t g c c t c c t g g g g c c t g g c c g c c c t g t g t g c c t a c a a g c g t g g c g t a a g	720
Qy	802	c a t g g g g t c t t g t c t c t c a c c a a g c c a c c t c g t t g c a t a t g g g t g g t g a t c	861
Db	721	c a t g g g g t c t t g t c t c t c a c a g c c a c c t c g t t g c a t a t g g g t g g t g g a t c	780
Qy	862	g t a t g t a c t t a c g g c a a g a g a c a a c a g t c c a c t g g a t g a c c c a c g t g	921
Db	781	g t a t g t a c t t a c g g c a a g a g a c a a c a g t c c c a c t g g a t g a c c c a c g t g	840
Qy	922	g c a t c c c t c g c c c a a t g c t g g g c c t t c t c t t t c a c t a c t c c c g a g g t c	981
Db	841	g c a t c c c t c g c c c a a t g c t g g g c c t t c g t c c t t c t a c g t a c t c c c g a g t c	900
Qy	982	t c c c a g t g a c c a a g t c c a g c c a g a a a g t a c c a g g g g a c a t g a c c c a c c g g	1041
Db	901	t c c c a g t g a c c a a g t c c a g c c a g a a a g t a c c a a g g g a c a t a t a c c a c c g g	960

Qy 913 cccagctgcccagctccgcccgaatgctggccttgcctctctctacgtcatc 972
Dy 895 cccaagctgcccagctccgcccgaatgctggccttgcctctctctacgtcatc 954
Qy 973 cccagagctctcccaggtgacaaagtcacagcccccagagcaaaagctaccaggggacatgtac 1032
Dy 955 cccagagctctcccaggtgacaaagtc--gccagagcaaaagctacca--gggacatgtac 1010
Qy 1033 cccacccgggg--cgtgggctatgagacatcctgaaagacagaaggtcagagcatgtt 1091
Dy 1011 cccaacccggggcgtgggctatgagacatcctgaaagacagaaggtcagagcatgtt 1070
Qy 1092 cgtgagaaacaagccttttccatgagagcgggtgacgctaaagagccgggtgtccacc 1151
Dy 1071 cgtgagaaacaagccttttccatgagagcgggtgacgctaaagagccgggtgtccacc 1130
Qy 1152 atacagcgggtacaaatgggagctgctgacagtggtgtaccagcccactgagatggccct 1211
Dy 1131 atacagcgggtacaaatgggagctgctgacagtggtgtaccagcccactgagatggccct 1190
Qy 1212 gatgcacaagttccgtccaaagagcttacgacatcatctccacggccacccgcaaa 1271
Dy 1191 gatgcacaagttccgtccaaagagcttacgacatcatctccacggccacccgcaaa 1250
Qy 1272 cagcaggtgatggcagtgccaaactgacccctgcggtggaagacatgtactcggccca 1331
Dy 1251 cagcaggtgatggcagtg--caactgcacccctgcggtggaagacatgtactcggccca 1309
Qy 1332 gagccacagcgggcccacacgcg--cgaaagacggcgaagaaactctcaggtctttagaaac 1390
Dy 1310 gagccacagcgggcccacacgcgaaagacggcgaagaaactctcaggtctttagaaac 1369
Qy 1391 cctcgtgtggactgagtcagcgtgagcggagagagcgggtcgatgtggagggccc 1450
Dy 1370 cctcgtgtggactgagtcagcgtgagcggagagcgggtcgatgtggagggccc 1429
Qy 1451 tgaggactggcccggcggcgaaggagactct--ccaggctctctctccctcggcagcccg 1509
Dy 1430 tgaggactggcccggcggcgaaggagactctccacggctctctccctcggcagcccg 1489
Qy 1510 caacatgtgcccagatgtggaaggcctcctctctgcagtggttgggtgggtgcat 1569
Dy 1490 caacatgtgcccagatgtggaaggcctcctctctgcagtggttgggtgggtgcat 1549
Qy 1570 ggggtcccacacactcctcagtggttggagtcagagagcgaacccagcctctgc 1629
Dy 1550 tgggtcccacacactcctcagtggttggagtcagagagcgaacccagcctctgc 1609
Qy 1630 caggatcacctcggggtcacactccagccaaatagttgttcggggtggtgggtgggca 1689
Dy 1610 caggatcacctcggggtcacactccagccaaatagttgttcggggtggtgggtgggca 1668
Qy 1690 gcgcctatgtttctctggagattcctgcaaacctcaagagactcccccagggcctcaggcct 1749
Dy 1669 gcgcctatgtttctctggagattcctgcaaacctcaagagactcccccagggcctcaggcct 1728
Qy 1750 ggaattgctcctctgtgagaaacaaggtgcctaaataaacatattctctgtttattaa 1808
Dy 1729 ggaattgctcctctgtgagaaacaaggtgcctaaataaacatattctctgtttattaa 1787

RESULT 12
AAZ32501

ID AAZ32501 standard; cDNA; 1326 BP.

XX AAZ32501;

AC AAZ32501;

XX AAZ32501;

DT 17-JAN-2000 (first entry)

DE G-protein coupled receptor AXOR4 polynucleotide sequence #1.

XX

,KW AXOR4; G-protein coupled receptor; RAIG1; pain; cancer; diabetes;

obesity; anorexia; bulimia; asthma; stroke; ulcer; allergy; infertility;
neurological disorder; schizophrenia; ss.
Homo sapiens.
Key Location/Qualifiers
CDS 1..1326
/*tag= a
/product= AXOR4
XX WO953054-A1.
XX 21-OCT-1999.
XX 07-APR-1999; 99WO-GB01067.
XX 08-APR-1998; 98GB-0007723.
XX (SMIK) SMITHKLINE BEECHAM PLC.
XX Medhurst A, Michalovich D, Pangalos M, Hill J;
WPI: 1999-620425/53.
P-PSDB; AAY49156.
XX New AXOR4 G-protein coupled receptor polypeptides and related nucleic
acid, for treating, preventing or diagnosing e.g. cancer -
Claim 6; Page 37-38; 51pp; English.
XX This is the G-protein coupled receptor AXOR4, polynucleotide sequence.
The AXOR4 polypeptide has homology with human putative G-protein coupled
receptor RAIG1. The AXOR4 polypeptide and polynucleotide sequences can be
used in the production of antibodies specific for AXOR4. The sequences
may also be used to screen for AXOR4 agonists or antagonists.
CC Alternatively the effect of a candidate agonist or antagonist compound on
the production of mRNA encoding AXOR4 may be detected using an ELISA
assay. Diseases or conditions arising from altered expression or activity
of AXOR4 may be diagnosed by detecting the AXOR4 protein in a sample from
a patient or detecting a mutation in the AXOR4 nucleotide sequence in the
genome of a patient. These diseases or conditions include pain, cancer,
diabetes, obesity, anorexia, bulimia, asthma, hypo- or hypertension,
stroke, ulcers, allergy, benign prostatic hypertrophy, migraine,
epilepsy, vomiting, psychosis, infertility, and neurological disorders
(e.g. anxiety, schizophrenia, depression, delirium, dementia,
neurodegeneration or severe mental retardation). Detection of the AXOR4
nucleotide sequence may be used for chromosome mapping and tissue
localization. The polynucleotide and polypeptide sequences may be
administered to patients as vaccines or as part of a gene therapy regime
respectively, to treat the diseases and conditions listed above.
XX Sequence 1326 BP; 232 A; 431 C; 386 G; 277 T; 0 other;

Query Match 72.88; Score 1324.4; DB 20; Length 1326;
Best Local Similarity 99.9%; Pred. No. 1.2e-292;
Matches 1325; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 82 atggccatccacaagccttggtgtagtgctggagactgctctctctctccaggg 141
Dy 1 atggccatccacaagccttggtgtagtgctggagactgctctctctctccaggg 60
Qy 142 gcttggccagggccatgtcccccggctgcagcagaagcctcaacccctgtactac 201
Dy 61 gcttggccagggccatgtcccccggctgcagcagaagcctcaacccctgtactac 120
Qy 202 aacctgtgtgaccgctctgtggcgctggtggcgatcgtctcctgtgagcgctggtggcgcg 261
Dy 121 aacctgtgtgaccgctctgtggcgctggtggcgatcgtctcctgtgagcgctggtggcgcg 180
Qy 262 atgtcaccacgtttgtgctcaccatcatcctggtgccagcctccctttgtgcagagc 321
Dy 181 atgtcaccacgtttgtgctcaccatcatcctggtgccagcctccctttgtgcagagc 240

Qy	83	tggccatccacaagacctgggtgatgtgctctgggactgcctctctctctgttctccagggg	142
Db	61	tggccatccacaagacctgggtgatgtgctctgggactgcctctctctgttctccagggg	120
Qy	143	cctgggcccaggccatgtgtccaccggctgagccaagcctcaacccctgtactaca	202
Db	121	cctgggcccaggccatgtgtccaccggctgagccaagcctcaacccctgtactaca	180
Qy	203	acctgtgtgaccgctctggggcgtggggcatcgtctctggagcgctggctggggcgagca	262
Db	181	acctgtgtgaccgctctggggcgtggggcatcgtctctggagcgctggctggggcgagca	240
Qy	263	ttgtcaacacgtttgtgtctcaacatcatctcgtgtggcagctcccctttgtgcaggaca	322
Db	241	ttgtcaacacgtttgtgtcaacatcatctcgtgtggcagctcccctttgtgcaggaca	300
Qy	323	ccaagaacaggacctgtgtgggaccaaggtattctctctggggaacctgggcctct	382
Db	301	ccaagaacaggacctgtgtgggaccaaggtattctctctggggaacctgggcctct	360
Qy	383	tctgcctcgtgtttgcctgtgtgtgaaoccgacctctccacctgtgcctctcggcgt	442
Db	361	tctgcctcgtgtttgcctgtgtgtgaaoccgacctctccacctgtgcctctcggcgt	420
Qy	443	tcctctttggggttctgttgtcccatctgtctctctgtgtggcggtcacgtctttggcc	502
Db	421	tcctctttggggttctgttgtcccatctgtctctctgtgtggcggtcacgtctttggcc	480
Qy	503	tcaacttcttggcccgaagaaaccaacgggcccggggctgggtgatcttcaactgtgctc	562
Db	481	tcaacttcttggcccgaagaaaccaacgggcccggggctgggtgatcttcaactgtgctc	540
Qy	563	tgtctgtacccctgttagaggtcatcataaatacagagtggctgatataccctgggtc	622
Db	541	tgtctgtacccctgttagaggtcatcataaatacagagtggctgatataccctgggtc	600
Qy	623	ggggcagtgggcagggcgccctcagggcaacagcagcgaggtcgggcgttgccctccc	682
Db	601	ggggcagtgggcagggcgccctcagggcaacagcagcgaggtcgggcgttgccctccc	660
Qy	683	cctgtgccatcgccaacatggaaatttgtatggcactcatatacgtcatgtctgtctgc	742
Db	661	cctgtgccatcgccaacatggaaatttgtatggcactcatatacgtcatgtctgtctgc	720
Qy	743	tgggtgctctccttgggggctggccgcctgtgtggccgctacaagcgctggcgtaagc	802
Db	721	tgggtgctctccttgggggctggccgcctgtgtggccgctacaagcgctggcgtaagc	780
Qy	803	atggggctttgtgtctctcaacacgaccactccgttgtccata tgggtgtgtggatcg	862
Db	781	atggggctttgtgtctctcaacacgaccactccgttgtccata tgggtgtgtggatcg	840
Qy	863	tcatgtatacttacggccaacagcaacacagtcaccaactcggatggatgaacccaactgg	922
Db	841	tcatgtatacttacggccaacagcaacacagtcaccaactcggatggatgaacccaactgg	900
Qy	923	ccatcgcccttcgcgcgaatggcctgggcttcgtctcttttaacgtatccccgaggtct	982
Db	901	ccatcgcccttcgcgcgaatggcctgggcttcgtctcttttaacgtatccccgaggtct	960
Qy	983	cccaggtgaccaagtccagcccagacaaagctacagggggacatgtacccccacccggg	1042
Db	961	cccaggtgaccaagtccagcccagacaaagctacagggggacatgtacccccacccggg	1020
Qy	1043	gcgtgggctatgagacatctctgaaagacagaagggctcagagcatgttctgtgagaaca	1102
Db	1021	gcgtgggctatgagacatctctgaaagacagaagggctcagagcatgttctgtgagaaca	1080
Qy	1103	agcccttttccatggatgacgggttgacgttaacagccgggtgtccacatacagcgggt	1162
Db	1081	agcccttttccatggatgacgggttgacgttaacagccgggtgtccacatacagcgggt	1140
Qy	1163	acaatgggcagctgctgaacagtggtgaacagcccaactgagatggccctgatgcacaag	1222

[illegible]

Db 1453 ctgaggacctgg-cccgggcaaggagactctccaggctcctccctccctggcaggccacg 1511

Qy 1510 caacatgtcccagatgtgaaggcctccctctctccagtgcttgggtgggtgtcat 1569

Db 1512 caacatgtcccagatgtgaaggcctccctctctccagtgcttgggtgggtgtcat 1571

Qy 1570 ggggtgtcccacactcctcagtggttggagtcgagagcgaacccacagctctctgc 1629

Db 1572 ggggtgtcccacactcctcagtggttggagtcgagagcgaacccacagctctctgc 1631

Qy 1630 caggatacctcggcgttcacatccagccaaatagtgttctcggggtggtcgtgggca 1689

Db 1632 caggatacctcggcgttcacatccagccaaatagtgttctcggggtggtcgtgggca 1691

Qy 1690 gcacctatgtttcttgagatcctcgaacctcaagagagacttcccaggcgtcaggcct 1749

Db 1692 gcacctatgtttcttgagatcctcgaacctcaagagagacttcccaggcgtcaggcct 1751

Qy 1750 ggaattgtcctctgtgaggaacaagggtgcctaataatacatttctgtttatt 1806

Db 1752 ggaattgtcctctgtgaggaacaagggtgcctaataatacatttctgtttatt 1808

RESULT 8

AAI59005

ID AAI59005 standard; cDNA; 1880 BP.

XX

AC AAI59005;

XX

DT 22-OCT-2001 (first entry)

XX

DE Human polynucleotide SEQ ID NO 1208.

XX

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;

KW peripheral nervous system; neuropathy; central nervous system; CNS;

KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;

KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;

KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;

KW leukaemia; ss.

XX

OS Homo sapiens.

XX

PN WO200153312-A1.

XX

PD 26-JUL-2001.

XX

PF 26-DEC-2000; 2000NO-US34263.

XX

PR 21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0552317.

PR 09-JUL-2000; 2000US-0598042.

PR 19-JUL-2000; 2000US-0620312.

PR 03-AUG-2000; 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.

PR 29-NOV-2000; 2000US-0727344.

XX

PA (HYSE-) HYSEQ INC.

XX

Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;

PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX

WPI; 2001-442253/47.

DR P-PSDB; AAM39849.

XX

Novel nucleic acids and polypeptides, useful for treating disorders

PT such as central nervous system injuries -

PT

XX

Claim 1: SEQ ID NO 1208; 10078pp; English.

PS

XX

The invention relates to human nucleic acids (AAI57798-AAI61369) and

CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,

CC

CC immunosuppressant and cytostatic activity. The polynucleotides are useful

CC in gene therapy. A composition containing a polypeptide or polynucleotide

CC of the invention may be used to treat diseases of the peripheral nervous

CC system, such as peripheral nervous injuries, peripheral neuropathy and

CC localised neuropathies and central nervous system diseases, such as

CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic

CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the

CC utilisation of the activities such as: Immune system suppression,

CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic

CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,

CC assays for receptor activity, arthritis and inflammation, leukaemias and

CC C.N.S disorders.

CC Note: The sequence data for this patent did not form part of the printed

CC specification.

XX

SQ Sequence 1880 BP; 350 A; 586 C; 551 G; 393 T; 0 other;

Query Match 98.1%; Score 1785; DB 22; Length 1880;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 1788; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 27 gtacgagtcggctcagcctggaggacccacagagcctggcctggagccaggatggc 86

Db 83 gtacgagtcggctcagcctggaggacccacagagcctggcctggagccaggatggc 142

Qy 87 catccacaaagccttggtgatgtgctgggactgcctctctctctgttccacaggggcctg 146

Db 143 catccacaaagccttggtgatgtgctgggactgcctctctctctgttccacaggggcctg 202

Qy 147 ggcacagggccatgtcccaccggctgcagcgaagcctcaacccctgtactacaacct 206

Db 203 ggcacagggccatgtcccaccggctgcagcgaagcctcaacccctgtactacaacct 262

Qy 207 gtgacacgctctggggcgtggggcatcgtccctggagcgcgtggcggcgccattgt 266

Db 263 gtgacacgctctggggcgtggggcatcgtccctggagcgcgtggcggcgccattgt 322

Qy 267 caccacgttctgtctcaccatcatctctgtgtggcagcctccctcttctgtgcagacaccaa 326

Db 323 caccacgttctgtctcaccatcatctctgtgtggcagcctccctcttctgtgcagacaccaa 382

Qy 327 gaaacgagcctcgtgggacccacagtgattctctctctctgtggacccctggcctctctg 386

Db 383 gaaacgagcctcgtgggacccacagtgattctctctctctgtgggacccctggcctctctg 442

Qy 387 cctcgtgttctgctgtgtgtgaagcccgacttctccacctgtgcctctcggcgcttct 446

Db 443 cctcgtgttctgctgtgtgtgaagcccgacttctccacctgtgcctctcggcgcttct 502

Qy 447 ctttggggttctgttcgcaatctgtctctctctctgttcgtggcggtcagcgtctttgccctcaa 506

Db 503 ctttggggttctgttcgcaatctgtctctctctgttcgtggcggtcagcgtctttgccctcaa 562

Qy 507 cttcctggcccggaagaacacacggcccggtggctgggtgatcttcaactgtgctctgct 566

Db 563 cttcctggcccggaagaacacacggcccggtggctgggtgatcttcaactgtgctctgct 622

Qy 567 gctgacctgtgtgagagtgatcatcatcaatcacagagtggtggtgatcatcaccttgctgggg 626

Db 623 gctgacctgtgtgagagtgatcatcatcaatcacagagtggtggtgatcatcaccttgctgggg 682

Qy 627 cagtgcgagggcgccctcaggggcaacagcagcagcgtggcgctggcgctcccccctg 686

Db 683 cagtgcgagggcgccctcaggggcaacagcagcagcgtggcgctggcgctcccccctg 742

Qy 687 tgcctatgcacacatgagactttgatgttgacatcatctctacgtcatctgctgctggg 746

Db 743 tgcctatgcacacatgagactttgatgttgacatcatctctacgtcatctgctgctggg 802

Qy 747 tgccttctcctggggcctggcccgccctgtgtggccgctacaagcgtggcgtaagcatgg 806

Db 803 tgccttctcctggggcctggcccgccctgtgtggccgctacaagcgtggcgtaagcatgg 862

PR 08-JUL-1999; 99JP-0194179.
 PR 11-JAN-2000; 2000JP-0118775.
 PR 02-MAY-2000; 2000JP-0183766.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Kawai Y, Sugiyama T, Hayashi K;
 XX WPI; 2001-093989/11.
 DR P-PSDB; AAB88359.
 XX
 XX Nucleic acids encoding secretory proteins/membrane proteins, useful in
 PT gene therapy or as candidate target molecules in drug development -
 XX
 XX Claim 1; SEQ ID 85; 609pp + CD ROM; English.
 CC This invention relates to nucleic acid sequences AAF93744 - AAF93916
 CC which encode human secretory or membrane proteins represented by
 CC AAB88317 - AAB88419. Included in the invention are primers
 CC AAF93917 - AAF94295 and AAF62232 - AAF62235 which are used to isolate the
 CC cDNA sequences of the invention. The invention also includes methods for
 CC the production of antibodies directed against the proteins, and cDNA
 CC sequences, which can be used in vaccines. The polynucleotide sequences
 CC can be used in gene therapy. The polynucleotide sequences and the
 CC proteins they encode may be used in the prevention, treatment and
 CC diagnosis of diseases associated with inappropriate secretory
 CC protein/membrane protein expression. The nucleic acids and complementary
 CC sequences may also be used as DNA probes in diagnostic assays
 CC (e.g. polymerase chain reactions (PCR)) to detect and quantitate the
 CC presence of similar nucleic acid sequences in samples. They may also be
 CC used to study the expression and function of secretory proteins/membrane
 CC polypeptides and their role in metabolism. The polypeptides may be used
 CC as antigens in the production of antibodies against them and in assays to
 CC identify modulators (agonists and antagonists) of expression and
 CC activity. The antibodies and antagonists may also be used as therapeutic
 CC agents to down regulate expression and activity. The antibodies may also
 CC be used as diagnostic agents for detecting the presence of the
 CC polypeptides in samples (e.g. by enzyme linked immunosorbent assay
 CC (ELISA). Examples of diseases which may be treated include rheumatoid
 CC arthritis and diabetes.
 XX
 XX Sequence 1808 BP; 321 A; 580 C; 531 G; 376 T; 0 other;

Db	373		acccctggcctcttctgcctcgtgttgcctgtgtgtggaagcccgactcttcacactgt	432
Qy	430	gcctctcggcgctctcttctgggttctgttcgccaatgccttctcttctgttggcggt	489	
Db	433	gcctctcggcgctctcttctgggttctgttcgccaatgccttctcttctgttggcggt	492	
Qy	490	cacgtctcttggccctcaacttctggcccggaagaacacacggcccggtggtggtatc	549	
Db	493	cacgtctcttggccctcaacttctggcccggaagaacacacggcccggtggtggtatc	552	
Qy	550	ttcactgtgctctgtctgacccctgtgtagaggtcatcatcaatacacagagtgtgctatc	609	
Db	553	ttcactgtgctctgtctgacccctgtgtagaggtcatcatcaatacacagagtgtgctatc	612	
Qy	610	atcacctctgttgcggcgagtgccgagggcgccctcagggcaacacagcagcagcgtg	669	
Db	613	atcacctctgttgcggcgagtgccgagggcgccctcagggcaacacagcagcagcgtg	672	
Qy	670	gccgtggcctccctctgcatcgccaacatgacatttgcattggcactcatcatcgtc	729	
Db	673	gccgtggcctccctctgcatcgccaacatgacatttgcattggcactcatcatcgtc	732	
Qy	730	atgctgctgtctgtgtgcttctctggtggcgcccgccctgtgtgcccgtacaag	789	
Db	733	atgctgctgtctgtgtgcttctctggtggcgcccgccctgtgtgcccgtacaag	792	
Qy	790	cgtctggcgttaagcatgggtcttctgtctctcaccacagccactcgttgcctatg	849	
Db	793	cgtctggcgttaagcatgggtcttctgtctctcaccacagccactcgttgcctatg	852	
Qy	850	gtggtgtggatcgtcatgtatacttacggcaacagcagcacacagtcaccactggat	909	
Db	853	gtggtgtggatcgtcatgtatacttacggcaacagcagcacacagtcaccactggat	912	
Qy	910	gacccacgctgtgccaatgcgcctgcgcgaatgcctggccttcgtctcttctatcgtc	969	
Db	913	gacccacgctgtgccaatgcgcctgcgcgaatgcctggccttcgtctcttctatcgtc	972	
Qy	970	atccccgaggtctcccaggtgacaaagtcacagccagcaaaagtcaccaggggacatg	1029	
Db	973	atccccgaggtctcccaggtgacaaagtcacagccagcaaaagtcaccaggggacatg	1032	
Qy	1030	taccccccggcggtggcgtatgagaccatcctgaaagcagagaggggtcagagcatg	1089	
Db	1033	taccccccggcggtggcgtatgagaccatcctgaaagcagagaggggtcagagcatg	1092	
Qy	1090	ttcgtgagaaacaaaggccttttccatggatgagccggttcagctaaagagccggtgtca	1149	
Db	1093	ttcgtgagaaacaaaggccttttccatggatgagccggttcagctaaagagccggtgtca	1152	
Qy	1150	ccatacagcgggtacaatggcgagctgtaccagtggtaccagcccaactgagatggcc	1209	
Db	1153	ccatacagcgggtacaatggcgagctgtaccagtggtaccagcccaactgagatggcc	1212	
Qy	1210	ctgatgcacaaagtcccgctccgaaggagcttacgacatcatcctcccacggccacgcgc	1269	
Db	1213	ctgatgcacaaagtcccgctccgaaggagcttacgacatcatcctcccacggccacgcgc	1272	
Qy	1270	aacagccaggtgtatggcgagtgccaactgcacctgcgggctgagagcatgtactcgcc	1329	
Db	1273	aacagccaggtgtatggcgagtgccaactgcacctgcgggctgagagcatgtactcgcc	1332	
Qy	1330	cagagccaccagcggccacaccccgcaaaagcagcgaagaactctcaggtcttttagaac	1389	
Db	1333	cagagccaccagcggccacaccccgcaaaagcagcgaagaactctcaggtcttttagaac	1392	
Qy	1390	ccctacgtgtggactgagtcagcgggtggcgagagagcggttcggatttggggagggcc	1449	
Db	1393	ccctacgtgtggactgagtcagcgggtggcgagagagcggttcggatttggggagggcc	1452	
Qy	1450	ctgaggacactggccccgggaaggagactctccaggtctctctcccctggcagggccag	1509	

Query Match 98.1%; Score 1785; DB 22; Length 1808;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1796; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy	10	ccctcaccagccgaaagtacagtcggtcagcctgaggggacccacacagagcctggc	69
Db	13	ccctcaccagccgaaagtacagtcggtcagcctgaggggacccacacagagcctggc	72
Qy	70	ctgggagccagatggcctacacaaagcctgtgtgtgctgctggagctctcttc	129
Db	73	ctgggagccagatggcctacacaaagcctgtgtgtgctgctggagctctcttc	132
Qy	130	ctgttccacagggcctggccagggcctatgccacccggtcagcgaaggcctcaac	189
Db	133	ctgttccacagggcctggccagggcctatgccacccggtcagcgaaggcctcaac	192
Qy	190	ccctgtactacaacctgtgtacgcctctggcgctgggcatcgtcctggagggcgtg	249
Db	193	ccctgtactacaacctgtgtacgcctctggcgctgggcatcgtcctggagggcgtg	252
Qy	250	gctggggcgccattgtcacacagtttgtctcaccatcatcctgggtggcagcctccc	309
Db	253	gctggggcgccattgtcacacagtttgtctcaccatcatcctgggtggcagcctccc	312
Qy	310	tttgtgagggacacaaagcagcctgctggggacccaggtattcttcttctggg	369
Db	313	tttgtgagggacacaaagcagcctgctggggacccaggtattcttcttctggg	372
Qy	370	accctgggccccttctgctcgtgtttgctgtgtgtggaagcccgacttctccactgt	429

Db 850 cgctggcgtgaagcatggggtcttctgctctcaccacagccacccctcgttgccatagg 909
QY 850 gtggtggtgagtcgtcatgtatctacgcgcaacaagcagcacacagtcaccactggat 909
Db 910 gtggtggtgagtcgtcatgtatctacgcgcaacaagcagcacacagtcaccactggat 969
QY 910 gaccccaagctggccatgcgcctcgcgcgaatgcctgggccccttcgtctctcttctacgc 969
Db 970 gaccccaagctggccatgcgcctcgcgcgaatgcctgggccccttcgtctcttctacgc 1029
QY 970 atcccgagctctccagctgaccaaagtcacgcccagcagcaaaagctaccaggggacatg 1029
Db 1030 atcccgagctctccagctgaccaaagtcacgcccagcagcaaaagctaccaggggacatg 1089
QY 1030 taccacacccggggtggtggtctatgagaccatctgaaagcagcagaggtcagagcatg 1089
Db 1090 taccacacccggggtggtggtctatgagaccatctgaaagcagcagaggtcagagcatg 1149
QY 1090 ttcgtggagaaacaggcctttccatggtgagccggttgagcctaagaagcccggtgtca 1149
Db 1150 ttcgtggagaaacaggcctttccatggtgagccggttgagcctaagaagcccggtgtca 1209
QY 1150 ccatacagcgggtacaatggcagctactacacagtggtaccagccacactgagatgcc 1209
Db 1210 ccatacagcgggtacaatggcagctactacacagtggtaccagccacactgagatgcc 1269
QY 1210 ctgatgcacaaagtccgtccgaaggagcttacgacatcatctccacagggccacccgcc 1269
Db 1270 ctgatgcacaaagtccgtccgaaggagcttacgacatcatctccacagggccacccgcc 1329
QY 1270 aacagccagtgatgggcagtgccacacacacacacacacacacacacacacacacac 1329
Db 1330 aacagccagtgatgggcagtgccacacacacacacacacacacacacacacacacac 1389
QY 1330 cagagccacacagcggccac 1389
Db 1390 cagagccacacagcggccac 1449
QY 1390 cctacgtgtggagctagtcagcgggtggcagagagcggcgtgatttgggagagcc 1449
Db 1450 cctacgtgtggagctagtcagcgggtggcagagagcggcgtgatttgggagagcc 1509
QY 1450 ctgaggacctggcccgggcaaggagctctccaggtctcctcctcctcctcctcctcctc 1509
Db 1510 ctgaggacctggcccgggcaaggagctctccaggtctcctcctcctcctcctcctcctc 1569
QY 1510 caacatgtgccagatgtgaaaggccctccctctctcagtgatttgggtggtgtcat 1569
Db 1570 caacatgtgccagatgtgaaaggccctccctctctcagtgatttgggtggtgtcat 1629
QY 1570 ggggtgcccccac 1629
Db 1630 ggggtgcccccac 1689
QY 1630 caggatcacctcgggtcacactccagcgaataagttctcgggtggtggtgggca 1689
Db 1690 caggatcacctcgggtcacactccagcgaataagttctcgggtggtggtgggca 1749
QY 1690 gcgcctatgttctctgagattctcgaacctcaaacctcaaacctcaaacctcaaacct 1749
Db 1750 gcgcctatgttctctgagattctcgaacctcaaacctcaaacctcaaacctcaaacct 1809
QY 1750 ggaattgtcctctctgagaaacagggtgcctcaataataacattctctgttatt 1806
Db 1810 ggaattgtcctctctgagaaacagggtgcctcaataataacattctctgttatt 1866

RESULT 6

AAX97994

ID AAX97994 standard; DNA; 1860 BP.

XX

AC AAX97994;

XX

DT 17-SEP-1999 (first entry)
XX Human secreted protein gene 79.
XX Human; secreted protein; cancer; tumour; developmental abnormality;
KW foetal deficiency; blood disorder; immune system disorder; inflammation;
KW autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;
KW schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;
KW atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;
KW digestive disorder; endocrine disorder; infection; AIDS; SS.
XX Homo sapiens.
OS WO9931117-A1.
XX PN 24-JUN-1999.
XX PD 17-DEC-1998; 98WO-US27059.
XX PF 19-DEC-1997; 97US-0068369.
XX PR 18-DEC-1997; 97US-0068006.
XX PR 18-DEC-1997; 97US-0068007.
XX PR 18-DEC-1997; 97US-0068008.
XX PR 18-DEC-1997; 97US-0068053.
XX PR 18-DEC-1997; 97US-0068054.
XX PR 18-DEC-1997; 97US-0068057.
XX PR 18-DEC-1997; 97US-0068064.
XX PR 18-DEC-1997; 97US-0070923.
XX PR 19-DEC-1997; 97US-0068169.
XX PR 19-DEC-1997; 97US-0068365.
XX PR 19-DEC-1997; 97US-0068367.
XX PR 19-DEC-1997; 97US-0068368.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Carter KC, Duan RD, Feng P, Ferrie AM, Florence C;
PI Florence K, Greene JM, Janat F, Kyaw H, Moore PA;
PI Ni J, Rosen CA, Ruben SM, Shi Y, Soppet DR, Wei Y;
PI Yu G;
XX WPI: 1999-418749/35.
XX P-PSDB; AAY36302.
XX New isolated human genes encoding secreted polypeptides
PS Claim 1; Page 319-320; 537pp; English.
XX AAX97916 to AAX98029 represent 110 isolated human secreted protein
CC genes. AAY36224 to AAY36727 represent the secreted proteins encoded by
CC the 110 human genes. The genes and their corresponding secreted
CC polypeptides are useful for preventing, treating or ameliorating medical
CC conditions, e.g. by protein or gene therapy. Also pathological conditions
CC can be diagnosed by determining the amount of the new polypeptides in a
CC sample or by determining the presence of mutations in the new genes.
CC Specific uses are described for each of the 110 genes, based on which
CC tissues they are most highly expressed in, and include developing
CC products for the diagnosis or treatment of cancer, tumours, developmental
CC abnormalities and foetal deficiencies, blood disorders, diseases of the
CC immune system, autoimmune diseases, inflammation, allergies, Alzheimer's
CC and cognitive disorders, schizophrenia, arthritis, asthma, psoriasis,
CC sepsis, skin disorders, atherosclerosis, diabetes, cardiovascular
CC disorders, kidney disorders, digestive/endocrine disorders, infections
CC and AIDS. The polypeptides are also useful for identifying their binding
CC partners. The sequences given in AAX97907 to AAX97915 and AAY36223 are
CC used in the exemplification of the present invention.
XX SQ Sequence 1860 BP; 348 A; 582 C; 543 G; 380 T; 7 other;

Query Match 98.2%; Score 1786.4; DB 20; Length 1860;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1798; Conservative 4; Mismatches 5; Indels 1; Gaps 1;

Claim 1; Page 688-689; 894pp; English.

The invention relates to polynucleotides encoding novel human proteins or their active domains. The polypeptides, polynucleotides and antibodies raised against the polypeptides are used in a method of treatment of a mammal and prevention of disorders caused by the aberrant protein expression or activity. The polypeptides can be used as molecular weight markers, food supplements, and in antibody production. The polypeptides are used to identify compounds which bind to the polypeptides. Polynucleotides of the invention are used as probes and primers, for sequencing, for chromosome or gene mapping, in the production of recombinant proteins, and in generating anti-sense DNA or RNA and in gene therapy. Polypeptides of the invention can be used to target drugs to a tumour, in assays to determine biological activity, to raise antibodies/elicite an immune response, to determine quantitative protein levels, as tissue markers, and to isolate receptors or ligands. Polypeptides of the invention may also be useful in treating platelet disorders, stem cell disorders, regenerating bone, cartilage, tendon, ligament and/or nerve tissue, wound healing, treating burns, promoting the proliferation, differentiation and survival of stem cells, as a contraceptive, treating osteoporosis and osteoarthritis, anaemia, Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral sclerosis, stroke, immune deficiencies resulting from bacterial, viral or fungal infection or from autoimmunity, cancer, allergy, asthma, graft-versus-host diseases, eczema, haemophilia, thrombosis, anti-inflammatory diseases, nervous system disorders, and infection. The present sequence encodes a protein of the invention.

Sequence 2027 BP; 370 A; 653 C; 607 G; 397 T; 0 other;

Query Match 98.8%; Score 1797.4; DB 22; Length 2027;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1801; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Db 761 ttcaactgtgctctgctgctgacccctggtagaggtcatcatcaatcacagagtggctgac 820
Qy 610 ataccctgttctcggggcagtgccgagggcgccctcagggcaacagcagcgagctgg 669
Db 821 ataccctgttctcggggcagtgccgagggcgccctcagggcaacagcagcgagctgg 880
Qy 670 gccgtggcctcccccctgctgaccccaatgagactttgtcatggtgacatctatcgtc 729
Db 881 gccgtggcctcccccctgctgaccccaatgagactttgtcatggtgacatctatcgtc 940
Qy 730 atgctgctgctgctgctgcttcttcttcttcttcttcttcttcttcttcttcttct 789
Db 941 atgctgctgctgctgctgcttcttcttcttcttcttcttcttcttcttcttcttct 1000
Qy 790 cgttggcgttaagcagtggttcttcttcttcttcttcttcttcttcttcttcttctt 849
Db 1001 cgttggcgttaagcagtggttcttcttcttcttcttcttcttcttcttcttcttctt 1060
Qy 850 gtggtgtggtatgctgctgcttcttcttcttcttcttcttcttcttcttcttcttct 909
Db 1061 gtggtgtggtatgctgcttcttcttcttcttcttcttcttcttcttcttcttctt 1120
Qy 910 gacccacgctggcagtcgcttcttcttcttcttcttcttcttcttcttcttcttctt 969
Db 1121 gacccacgctggcagtcgcttcttcttcttcttcttcttcttcttcttcttcttctt 1180
Qy 970 atccccaggttctccaggtgacaaagtcacagccagagcaaaagtcacaggggacatg 1029
Db 1181 atccccaggttctccaggtgacaaagtcacagccagagcaaaagtcacaggggacatg 1240
Qy 1030 taccaccccccggcgctgggtatgagaccatctctgaaagagcagaaggttcaagagcatg 1089
Db 1241 taccaccccccggcgctgggtatgagaccatctctgaaagagcagaaggttcaagagcatg 1300
Qy 1090 ttctgtgagaacaaagccttttcttcttcttcttcttcttcttcttcttcttcttct 1149
Db 1301 ttctgtgagaacaaagccttttcttcttcttcttcttcttcttcttcttcttcttct 1360
Qy 1150 ccatacagcgggtacaaatggcgagctgtgaccagttgttaccagccctactgagatggcc 1209
Db 1361 ccatacagcgggtacaaatggcgagctgtgaccagttgttaccagccctactgagatggcc 1420
Qy 1210 ctgtatgacaaagtctcgtccgaagagcttaccagacatctctccacgggacacaccc 1269
Db 1421 ctgtatgacaaagtctcgtccgaagagcttaccagacatctctccacgggacacaccc 1480
Qy 1270 aacagccagctgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 1329
Db 1481 aacagccagctgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 1540
Qy 1330 cagagccacagcggcggccacacggcgaagagcggcaagaaactctcaagctcttttaga 1389
Db 1541 cagagccacagcggcggccacacggcgaagagcggcaagaaactctcaagctcttttaga 1600
Qy 1390 cctacagtggtggactgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 1449
Db 1601 cctacagtggtggactgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 1660
Qy 1450 ctgagagacctggccccggcgcaagggactctccaggtctctctccctctccctggcag 1509
Db 1661 ctgagagacctggccccggcgcaagggactctccaggtctctctccctctccctggcag 1720
Qy 1510 caacatgtgccccagatgtggaaggcctctctctctctctctctctctctctctctctct 1569
Db 1721 caacatgtgccccagatgtggaaggcctctctctctctctctctctctctctctctctct 1780
Qy 1570 ggggtgtcccaac 1629
Db 1781 ggggtgtcccaac 1840
Qy 1630 caggatcacctcgggtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 1689
Db 1841 caggatcacctcgggtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 1900

```

|||||
Db 255 gctggggcgccattgctaccacgcttctgtctcaccatcactctggtggccagcctcccc 314
Qy 310 ttgtgtaggacacaaagacgagcctgtctgggacccaggtattctctcttctggg 369
Db 315 ttgtgtaggacacaaagacgagcctgtctgggacccaggtattctctcttctggg 374
Qy 370 accctggccctctctgctcgtgtttccctgtgtgtggaagcccgacttctccactgt 429
Db 375 accctggccctctctgctcgtgtttccctgtgtgtggaagcccgacttctccactgt 434
Qy 430 gacctgagcgtctctcttggggtctgttcgcacatctgtcttcttctgtctggcgct 489
Db 435 gacctgagcgtctctcttggggtctgttcgcacatctgtcttcttctgtctggcgct 494
Qy 490 cagctctttgccctcaactctctggcccggaagaaacccagggcccccgggctgggtgatc 549
Db 495 cagctctttgccctcaactctctggcccggaagaaacccagggcccccgggctgggtgatc 554
Qy 550 ttcaactgtgctctgctgacccctgttagaggtcaatcaatcacagagtgctgatc 609
Db 555 ttcaactgtgctctgctgacccctgttagaggtcaatcaatcacagagtgctgatc 614
Qy 610 atcaacctggttcggggcagtgagggcgccctcagggcacaacagcagcgagcgtgg 669
Db 615 atcaacctggttcggggcagtgagggcgccctcagggcacaacagcagcgagcgtgg 674
Qy 670 gccgtggccctccccctgtgccatcgccaacatggactttgcatggcaactatctacgtc 729
Db 675 gccgtggccctccccctgtgccatcgccaacatggactttgcatggcaactatctacgtc 734
Qy 730 atgctactgctggtgacctctctcggggcgctggcccgccctgtgtggccgctacaag 789
Db 735 atgctactgctggtgacctctctcggggcgctggcccgccctgtgtggccgctacaag 794
Qy 790 cgctggcgtgaagcatgggtctttgtctcctccacacagccacccctccgttggccatgg 849
Db 795 cgctggcgtgaagcatgggtctttgtctcctccacacagccacccctccgttggccatgg 854
Qy 850 gtggtgtgtagctcatgtatacttaacggcaacaaacagcaacacagctccacctggatg 909
Db 855 gtggtgtgtagctcatgtatacttaacggcaacaaacagcaacacagctccacctggat 914
Qy 910 gacccacgctggccatcgccctcgccgcaatggctgggctctcctctctctacgtc 969
Db 915 gacccacgctggccatcgccctcgccgcaatggctgggctctcctctctctacgtc 974
Qy 970 atccccaggtctcccaggtgaccaagttccagcccagagcaaaagctaccagggggacatg 1029
Db 975 atccccaggtctcccaggtgaccaagttccagcccagagcaaaagctaccagggggacatg 1034
Qy 1030 taccceacccggggcggtgggtatgagaccatctctgaaagagcagaaggtcagagcatg 1089
Db 1035 taccceacccggggcggtgggtatgagaccatctctgaaagagcagaaggtcagagcatg 1094
Qy 1090 ttcgtggagacaagcgttttccatgtagtgagcgggttcagctaagagccgggtgtca 1149
Db 1095 ttcgtggagacaagcgttttccatgtagtgagcgggttcagctaagagccgggtgtca 1154
Qy 1150 ccatcacagcgggtacaatggtggcagctgctgaccagtgctgtaccagccactgagatggcc 1209
Db 1155 ccatcacagcgggtacaatggtggcagctgctgaccagtgctgtaccagccactgagatggcc 1214
Qy 1210 ctgagtacaaaagttccgtccgaagagcttacgacatcatctccacagggccacccg 1269
Db 1215 ctgagtacaaaagttccgtccgaagagcttacgacatcatctccacagggccacccg 1274
Qy 1270 aacagccaggtgtatgggcagtgccaactcgacctgcgggcggtgaagacatgtactcgcc 1329
Db 1275 aacagccaggtgtatgggcagtgccaactcgacctgcgggcggtgaagacatgtactcgcc 1334
Qy 1330 cagagccaccagggcgccacaccccgcaagacggcaagaactctcaggtctttagaac 1389
|||||
```

```

Db 1335 cagagccaccagggcgccacacggccgaaagacggcgaaagactctcaggtctttagaac 1394
Qy 1390 ccctacatgtggaactagtcacggtggtcgagagagcggtcggtatttgggagggcc 1449
Db 1395 ccctacatgtggtggactgagtcacggtggtcgagagagcggtcggtatttgggagggcc 1454
Qy 1450 ctgaggaactggccccgggcaagggactctccaggtctctctctccctggcagggccag 1509
Db 1455 ctgaggaactggccccgggcaagggactctccaggtctctctctccctggcagggccag 1514
Qy 1510 caacatgtgccccagatgtggaagggcctccctctctcagtgatttgggtgggtgtcat 1569
Db 1515 caacatgtgccccagatgtggaagggcctccctctctcagtgatttgggtgggtgtcat 1574
Qy 1570 ggggtgccccaccactctcagtgatttggagtcgagagcgaacccagcctctgc 1629
Db 1575 ggggtgccccaccactctcagtgatttggagtcgagagcgaacccagcctctgc 1634
Qy 1630 cagatcacctcggcgttcacactccagccaaaatagtgattctcgggtggtggctgggca 1689
Db 1635 caggatcacctcggcgttcacactccagccaaaatagtgattctcgggtggtggctgggca 1694
Qy 1690 ggcctatgtttctctgagatctcctcgaacctcaagagacttccagggcgtcaggcct 1749
Db 1695 ggcctatgtttctctgagatctcctcgaacctcaagagacttccagggcgtcaggcct 1754
Qy 1750 ggaattgtctctctgtgaggaacaagggtgcttaataataacatttctgtttttaaac 1809
Db 1755 ggaattgtctctctgtgaggaacaagggtgcttaataataacatttctgtttttaaaa 1814
Qy 1810 tcttaaaaaa 1819
Db 1815 aaaaaaaaaa 1824
|||||
RESULT 4
AAS22707
ID AAS22707 standard; cDNA; 2027 BP.
XX
AC AAS22707;
XX
DT 24-OCT-2001 (first entry)
XX
DE Human cDNA encoding a novel human protein #273.
XX
KW Human; novel protein; ss; Antianemic; osteopathic; antiinflammatory;
immunomodulatory; cytostatic; neuroprotective; vulnery; nootropic;
anticonvulsant; antiallergic; cerebroprotective; antifungal; antiviral;
antibacterial; antiallergic; dermatological; haemostatic; antiasthmatic;
thrombolytic; immunogen; antibody; gene therapy; neurological disorder;
Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;
tissue regeneration; immune disorder.
XX
OS Homo sapiens.
XX
PN WO200155437-A2.
XX
PD 02-AUG-2001.
XX
PF 25-JAN-2001; 2001WO-US02623.
XX
PR 25-JAN-2000; 2000US-0491404.
XX
(HYSE-) HYSEQ INC.
XX
Tang YT, Liu C, Drmanac RT;
XX
WPI: 2001-451939/48.
XX
P-PSDB; AAU14402.
XX
Isolated polypeptides useful for treating anti-inflammatory diseases,
PT nervous system disorders, and for regenerating bone and cartilage -
XX
```

```
QY 970 atcccagaggtctccaggtagaccaggtccagccagagcaaaagctaccaggggagcatg 1029
Db 1238 atcccagaggtctccaggtagaccaggtccagccagagcaaaagctaccaggggagcatg 1297
QY 1030 taccccacccggggcgtggcctatgagaccatctctaaagagcagaagggctcagagcatg 1089
Db 1298 taccccacccggggcgtggcctatgagaccatctctaaagagcagaagggctcagagcatg 1357
QY 1090 ttctgtgagaaacaggccttttccatggatgagccgggttgacagtaagagccgggtgtca 1149
Db 1358 ttctgtgagaaacaggccttttccatggatgagccgggttgacagtaagagccgggtgtca 1417
QY 1150 ccatacagcgggttacaatggcagctgctgaccagtggtgtaccagcccaactgagatggcc 1209
Db 1418 ccatacagcgggttacaatggcagctgctgaccagtggtgtaccagcccaactgagatggcc 1477
QY 1210 ctgatgcacaaatctccgtccgaagagcttaccagatcatctctccacgggcccacgcgc 1269
Db 1478 ctgatgcacaaatctccgtccgaagagcttaccagatcatctctccacgggcccacgcgc 1537
QY 1270 aacagccaggtgatggcagtgccaactgcacccctcgccggtggaagacatgtactcggcc 1329
Db 1538 aacagccaggtgatggcagtgccaactgcacccctcgccggtggaagacatgtactcggcc 1597
QY 1330 cagagccaccagggccacacccgcgaagacggcaagaaactctcaggtcttttagaaac 1389
Db 1598 cagagccaccagggccacacccgcgaagacggcaagaaactctcaggtcttttagaaac 1657
QY 1390 cctacgtgtggagactgagtcagcgggtggagagagagcggctcgagatttggggagggcc 1449
Db 1658 cctacgtgtggagactgagtcagcgggtggagagagagcggctcgagatttggggagggcc 1717
QY 1450 ctgaggaacctggcccgggcgaaggactctccaggtctctctccctccctggcagggccag 1509
Db 1718 ctgaggaacctggcccgggcgaaggactctccaggtctctctccctccctggcagggccag 1777
QY 1510 caacatgtgcccagatgtggaaggccctccctctctgcagtggttgggtgggtgtcat 1569
Db 1778 caacatgtgcccagatgtggaaggccctccctctctgcagtggttgggtgggtgtcat 1837
QY 1570 ggggtgtcccccacacactctcagttgttggagtcgagagagcaacccagcctctgc 1629
Db 1838 ggggtgtcccccacacactctcagttgttggagtcgagagagcaacccagcctctgc 1897
QY 1630 caggatcacctcgcggtcacactccagccaatagttcttcggggtggtggtgggca 1689
Db 1998 caggatcacctcgcggtcacactccagccaatagttcttcggggtggtggtgggca 1957
QY 1690 gcgcctatgttctctggagattctctgcaacctcaagagagacttcccaggcgtcagggcct 1749
Db 1958 gcgcctatgttctctggagattctctgcaacctcaagagagacttcccaggcgtcagggcct 2017
QY 1750 ggaactctgctctctgtgagaaacagggtgccttaataataattctctttattaac 1809
Db 2018 ggaactctgctctctgtgagaaacagggtgccttaataataattctctttattaaa 2077
QY 1810 tcttaa 1816
Db 2078 aaaaaa 2084

RESULT 3
AAZ43798
ID AAZ43798 standard; cDNA; 1936 BP.
XX
AC AAZ43798;
XX
DT 10-MAR-2000 (first entry)
XX
DE Human fetal brain cDNA clone vc36_1.
XX
KW Human; secreted protein; treatment; nutritional activity; cytokine;
,KW cell proliferation; cell differentiation; hematopoiesis regulation;
```

```
KW tissue growth; activin; inhibin; chemotactic; chemokinetic; hemostatic;
KW thrombolytic; anti-inflammatory; invasion suppressor; tumor inhibition;
KW gene therapy; ds.
XX Homo sapiens.
XX WO955721-A1.
PN 04-NOV-1999.
XX
XX 23-APR-1999; 99WO-US08504.
XX
XX 24-APR-1998; 98US-0082904.
PR 11-JUN-1998; 98US-0088994.
PR 12-JUN-1998; 98US-0089278.
PR 02-JUL-1998; 98US-0091647.
PR 24-AUG-1998; 98US-0097639.
PR 22-APR-1999; 99US-0097639.
XX
XX (ALPH-) ALPHAGENE INC.
XX Valenzuela D, Yuan O, Hoffman H, Hall J, Rapiejko P;
PI WPI; 2000-052801/04.
XX P-PSDB; AAY50932, AAY50933.
DR
XX New polynucleotides encoding secreted human proteins, derived from
PT human fetal brain, adult skin, adult brain, adult heart, adult thymus
PT and adult aorta cDNA libraries.
XX
XX Claim 52a; Page 245-246; 282pp; English.
XX
XX This invention describes novel human secreted proteins which are encoded
CC by polynucleotides obtained from fetal brain, adult skin, adult brain,
CC adult heart, adult thymus and adult aorta cDNA libraries. The
CC polynucleotides and proteins are predicted to have biological activities
CC which would make them suitable for treating, preventing or ameliorating
CC medical conditions in humans and animals, although no supporting data
CC is given. Suggested activities include nutritional activity, cytokine
CC and cell proliferation/differentiation activity, immune stimulating
CC (e.g. as vaccines) or suppressing activity, hematopoiesis regulating
CC activity, tissue growth activity, activin/inhibin activity,
CC chemotactic/chemokinetic activity, hemostatic and thrombolytic activity,
CC receptor/ligand activity, anti-inflammatory activity, cadherin/tumor
CC invasion suppressor activity, and tumor inhibition activity. The
CC polynucleotides are also stated to be useful for gene therapy.
CC AAZ4377-243808 represent the polynucleotides described in the invention
CC which encode the proteins represented in AAY50905-150947.
XX
XX Sequence 1936 BP; 449 A; 581 C; 532 G; 374 T; 0 other;
SQ
Query Match 98.9%; Score 1798.8; DB 21; Length 1936;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1803; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 10 ccctcacccagccggaagtagcagctcagcctgagggagggaccacacagagcctggc 69
Db 15 ccctcacccagccggaagtagcagctcagcctgagggagggaccacacagagcctggc 74
QY 70 ctgggagccaggatggccatccacaaagccttggtgtagtgctggagactgcctctcttc 129
Db 75 ctgggagccaggatggccatccacaaagccttggtgtagtgctggagactgcctctcttc 134
QY 130 ctgttcccaggggcctggcccagggccatgtcccacccgctgcagccagggcctcaac 189
Db 135 ctgttcccaggggcctggcccagggccatgtcccacccgctgcagccagggcctcaac 194
QY 190 cccctgtactacaacctgtgtgacctcctggggcgctggggcgatcgtcctggagggcgtg 249
Db 195 cccctgtactacaacctgtgtgacctcctggggcgctggggcgatcgtcctggagggcgtg 254
QY 250 gctggggcgggcattgtcaccaogtttgtgtcaccatcatcctgtgtggcagcctcccc 309
```

QY 1741 ctgagcgtgactgtctctctgtgaggaacaagggtgcctcaataataacattttctgc 1800
|||||
Db 1741 ctgagcgtgactgtctctctgtgaggaacaagggtgcctcaataataacattttctgc 1800

QY 1801 ttattataactcttaaaaaa 1819
|||||
Db 1801 ttattataactcttaaaaaa 1819

RESULT 2
AAS22471
ID AAS22471 standard; cDNA; 2084 BP.
XX
AC AAS22471;
XX
XX 24-OCT-2001 (first entry)
XX
DE Human cDNA encoding a novel human protein #37.
KW Human; novel protein; ss; Antianaemic; osteopathic; antiinflammatory;
KW immunomodulatory; cytostatic; neuroprotective; vulnerary; nootropic;
KW anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral;
KW antibacterial; antiallergic; dermatological; haemostatic; antiasthmatic;
KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;
KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;
KW tissue regeneration; immune disorder.
XX
OS Homo sapiens.
XX
XX WO200155437-A2.
XX
XX 02-AUG-2001.
XX
XX 25-JAN-2001; 2001WO-US02623.
XX
XX 25-JAN-2000; 2000US-0491404.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX
XX WPI; 2001-451939/48.
XX
XX P-PSDB; AAU14166.
XX
XX Isolated polypeptides useful for treating anti-inflammatory diseases,
XX nervous system disorders, and for regenerating bone and cartilage -
XX
XX Claim 1; Page 226-228; 894pp; English.
XX
XX The invention relates to polynucleotides encoding novel human
XX proteins or their active domains. The polypeptides, polynucleotides and
XX antibodies raised against the polypeptides are used in a method of
XX treatment of a mammal and prevention of disorders caused by the aberrant
XX protein expression or activity. The polypeptides can be used as
XX molecular weight markers, food supplements, and in antibody production.
XX The polypeptides are used to identify compounds which bind to the
XX polypeptides. Polynucleotides of the invention are used as probes and
XX primers, for sequencing, for chromosome or gene mapping, in the
XX production of recombinant proteins, and in generating anti-sense DNA or
XX RNA and in gene therapy. Polypeptides of the invention can be used to
XX target drugs to a tumour, in assays to determine biological activity, to
XX raise antibodies/elicit an immune response, to determine quantitative
XX protein levels, as tissue markers, and to isolate receptors or ligands.
XX Polypeptides of the invention may also be useful in treating platelet
XX disorders, stem cell disorders, regenerating bone, cartilage, tendon,
XX ligament and/or nerve tissue, wound healing, treating burns, promoting
XX the proliferation, differentiation and survival of stem cells, as a
XX contraceptive, treating osteoporosis and osteoarthritis, anaemia, as
XX Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral
XX sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
XX fungal infection or from autoimmunity, cancer, allergy, asthma,
XX graft-versus-host disease, eczema, haemophilia, thrombosis,
XX anti-inflammatory diseases, nervous system disorders, and infection.

CC The present sequence encodes a protein of the invention.
XX
SQ Sequence 2084 BP; 378 A; 673 C; 625 G; 408 T; 0 other;

Query Match 98.9%; Score 1799; DB 22; Length 2084;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1802; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 10 cctcaccagccggaaaagtacagtggtcagctcagctggagggagcccaacagagcctggc 69
|||||
Db 278 cctcaccagccggaaaagtacagtggtcagctcagctggagggagcccaacagagcctggc 337

QY 70 ctgggagccagagatgcccatcccaaaagccttgatgtgcctggagctcctctcttc 129
|||||
Db 338 ctgggagccagagatgcccatcccaaaagccttgatgtgcctggagctcctctcttc 397

QY 130 ctgttcccaggggcctgggcccagggccatgtcccacccggctgcagcccaagccctcaac 189
|||||
Db 398 ctgttcccaggggcctgggcccagggccatgtcccacccggctgcagcccaagccctcaac 457

QY 190 ccctctactacaaacctgtgtgacccgtctcctggggcgtggggcatcgtcctggagggcgtg 249
|||||
Db 458 ccctctactacaaacctgtgtgacccgtctcctggggcgtggggcatcgtcctggagggcgtg 517

QY 250 gctggggcgggaattgtcaccacagtgtgtcaccacatcctcctggtggccagcctcccc 309
|||||
Db 518 gctggggcgggaattgtcaccacagtgtgtcaccacatcctcctggtggccagcctcccc 577

QY 310 ttgtgcaggagacaccagaaacggagcctgctggggacccaggtattcttctctctgggg 369
|||||
Db 578 ttgtgcaggagacaccagaaacggagcctgctggggacccaggtattcttctctctgggg 637

QY 370 accctgggaccttctgacctgtgtgttgcctgtgtgtgagcccgacctctccacctgt 429
|||||
Db 638 accctgggaccttctgacctgtgtgttgcctgtgtgtgagcccgacctctccacctgt 697

QY 430 gcctctggcgcttctcttctgggggtctgttcgcacatcctgcttctctctctctggggct 489
|||||
Db 698 gcctctggcgcttctcttctgggggtctgttcgcacatcctgcttctctctctctggggct 757

QY 490 cagctcttgcctcaactctcctggccggagaaacacacggccggcgctgggtgctgac 549
|||||
Db 758 cagctcttgcctcaactctcctggccggagaaacacacggccggcgctgggtgctgac 817

QY 550 ttcaactgtgctcgtcgtgacctggttagaggttcaatcaatacagagtggtgctgac 609
|||||
Db 818 ttcaactgtgctcgtcgtgacctggttagaggttcaatcaatacagagtggtgctgac 877

QY 610 atcacctggttcggggcagtgagggcgccctcaggggcaacagcagcagcagctgg 669
|||||
Db 878 atcacctggttcggggcagtgagggcgccctcaggggcaacagcagcagcagcagctgg 937

QY 670 gcctggcctccctcctgctgcacatcgcaacatggacctttgcatggcacctcctacgtc 729
|||||
Db 938 gcctggcctccctcctgctgcacatcgcaacatggacctttgcatggcacctcctacgtc 997

QY 730 atgctgctcgtcgtgggtgccttctcctggggcctggcccgccctgtgtggccctacaag 789
|||||
Db 998 atgctgctcgtcgtgggtgccttctcctggggcctggcccgccctgtgtggccctacaag 1057

QY 790 cgtggcgtaagcatggggtcttctgctcctcaccacagcaccctcctggtgcacatgg 849
|||||
Db 1058 cgtggcgtaagcatggggtcttctgctcctcaccacagcaccctcctggtgcacatgg 1117

QY 850 gtggtgtgatcgtcgtatcttacggcaacaagcagcacacagctccacctgggat 909
|||||
Db 1118 gtggtgtgatcgtcgtatcttacggcaacaagcagcacacagctccacctgggat 1177

QY 910 gacccacacgtggccatcgccctcgccgcaatgctggcgcttctctctctctacgtc 969
|||||
Db 1178 gacccacacgtggccatcgccctcgccgcaatgctggcgcttctctctctctacgtc 1237

DR P-PSDB; AAY57283.

XX Human G protein coupled protein receptor peptides useful for the
PT prevention, diagnosis and treatment of cell proliferative, neurological
PT and immune disorders -

XX Claim 9; Page 67; 71pp; English.

XX The invention provides human G protein coupled protein receptor (HGPRP)
CC polypeptides and polynucleotides encoding them. The polypeptides can be
CC produced by standard recombinant methodology. The polynucleotides and
CC polypeptides may be used in the prevention, treatment and diagnosis of
CC diseases associated with their inappropriate expression. Diseases that
CC can be treated are cell proliferative disorders (e.g. arteriosclerosis,
CC atherosclerosis and hepatitis), cancers (e.g. leukemia, melanomas and
CC adenocarcinoma), immune disorders (e.g. anemia, asthma and Crohn's
CC disease) and neurological disorders (e.g. epilepsy, Alzheimer's disease
CC and Parkinson's disease). The anti-HGPRP antibodies may also be used as
CC diagnostic agents for detecting the presence of HGPRP polypeptides in
CC samples (e.g. by enzyme linked immunosorbant assay (ELISA)). Sequences
CC AA290521-526 represent cDNA fragments encoding the HGPRP polypeptides
CC (AAY57283-286).

XX Sequence 1824 BP; 331 A; 581 C; 535 G; 377 T; 0 other;

Query Match 100.0%; Score 1819; DB 21; Length 1824;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1819; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	cggtcagagccctaccagccggaagtacagtagtcggtcagctcagctgagggaccacaacca	60
Db	1	cggtcagagccctaccagccggaagtacagtagtcggtcagctcagctgagggaccacaacca	60
Qy	61	gagctggctgggagcagatggccatccacaaagccttggtgatgtcgtggactg	120
Db	61	gagctggcctgggagcagatggccatccacaaagccttggtgatgtcgtggactg	120
Qy	121	cctctctctctccagggcctggccagggccatgtcccacccggctgcagccaa	180
Db	121	cctctctctctccagggcctggccagggccatgtcccacccggctgcagccaa	180
Qy	181	ggctcaccacccctgtactacaacctgtgtgaccgctctctgggctgggcatcgtctg	240
Db	181	ggctcaccacccctgtactacaacctgtgtgaccgctctctgggctgggcatcgtctg	240
Qy	241	gagccgtggtggcgggcagctgtacacacgtttgtgtccacatcatcctggtggcc	300
Db	241	gagccgtggtggcgggcagctgtacacacgtttgtgtccacatcatcctggtggcc	300
Qy	301	agctccctcttgtcagggacaccaaagacgagcctgctgggacccaggtattcttc	360
Db	301	agctccctcttgtcagggacaccaaagacgagcctgctgggacccaggtattcttc	360
Qy	361	cttctggggaccctgggacctcttctgctctgttttgcctgtgtggtgaagcccgacttc	420
Db	361	cttctggggaccctgggacctcttctgctctgttttgcctgtgtggtgaagcccgacttc	420
Qy	421	tcacactgtgctctcggcgctctcttctgtgggttctgttcgcacatgcttcttctgt	480
Db	421	tcacactgtgctctcggcgctctcttctgtgggttctgttcgcacatgcttcttctgt	480
Qy	481	ctggcggtcacgtcttggcctcaactctctggccgggaagacacacggcccggggc	540
Db	481	ctggcggtcacgtcttggcctcaactctctggccgggaagacacacggcccggggc	540
Qy	541	tgggtgatcttcaactgtggtctgtgctgacctgggtagaggtcatcatcaatcacagag	600
Db	541	tgggtgatcttcaactgtggtctgtgctgacctgggtagaggtcatcatcaatcacagag	600
Qy	601	tggctgatcatcaacctggttcggggcagtagggcagggcgccctcagggcaacagcagc	660
Db	601	tggctgatcatcaacctggttcggggcagtagggcagggcgccctcagggcaacagcagc	660

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2002, 15:34:27 ; Search time 292.25 seconds
(without alignments)
10686.287 Million cell updates/sec

Title: US-09-895-686-7
Perfect score: 1819
Sequence: 1 cggctcgagccctaccagc.....ctttatttaactcttaaaaaa 1819

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues
Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match %	Length DB ID	Description
1	1819	100.0	1824 21	AAZ90521 Human GPCR protein
2	1799	98.9	2084 22	AAS22471 Human cDNA encodin
3	1798.8	98.9	1936 21	AZ43798 Human fetal brain
4	1797.4	98.8	2027 22	AAS22707 Human cDNA encodin
5	1797	98.8	1866 22	AAF94491 Human hydrophobic
6	1786.4	98.2	1860 20	AAAX9794 Human secreted pro
7	1785	98.1	1808 22	AAF93786 Human cDNA encodin
8	1785	98.1	1880 22	AAI59005 Human polynucleoti
9	1784.4	98.1	1794 22	AAI60791 Human polynucleoti

10	1783.8	98.1	2406	21	AAZ29292	Human Metabotropic
11	1630.4	89.6	1790	20	AAZ32502	G-protein coupled
12	1324.4	72.8	1326	20	AAZ32501	G-protein coupled
13	1323	72.7	1323	22	AAF94481	Human hydrophobic
14	1310.2	72.0	1467	21	AAZ29294	Polynucleotide pro
15	682.2	37.5	1034	20	AAZ98029	Human secreted pro
16	640.8	35.2	684	22	AAH03345	Human cDNA clone (
17	624	34.3	891	21	AAZ29293	Mouse Metabotropic
18	587.8	32.3	946	22	AAI99248	Human excretory re
19	587.8	32.3	946	22	AAI63598	Human kidney relat
20	575.8	31.7	613	22	AAF93982	Primer specific fo
21	271.8	14.9	3362	21	AAZ77294	Human ORFX ORF2849
22	270.2	14.9	1212	20	AAZ20297	Human G-protein co
23	270.2	14.9	1520	21	AAZ90525	Human GPCR protein
24	270.2	14.9	2389	22	AAK51748	Human polynucleoti
25	270.2	14.9	2749	22	AAK94479	Human full-length
26	270.2	14.9	3852	20	AAZ20298	Human G-protein co
C 27	255	14.0	540	22	AAH08883	Human cDNA clone (
C 28	228.4	12.6	251	22	AAF94129	Primer specific fo
C 29	207.8	11.4	566	22	AAI99247	Human excretory re
30	207.8	11.4	566	22	AAI63597	Human kidney relat
31	202.8	11.1	207	20	AAZ40620	Human secreted pro
C 32	202.4	11.1	1000	22	ABA48935	Human breast cell
C 33	202.4	11.1	1000	22	ABA66850	Human foetal liver
C 34	202.4	11.1	1000	22	ABA33918	Probe #12384 for g
C 35	202.4	11.1	1000	22	AAK15285	Human brain expres
C 36	202.4	11.1	1000	22	AAK41008	Human bone marrow
C 37	202.4	11.1	1000	22	AAI21778	Probe #11711 for g
C 38	202.4	11.1	1000	22	AAI47061	Probe #15747 used
C 39	202.4	11.1	1000	22	AAI07463	Probe #7454 used t
C 40	201.8	11.1	205	21	AAC00605	Human secreted pro
41	201.4	11.1	1619	22	AAF58615	Human RECAP polynu
42	201.4	11.1	2446	22	AAH14688	Human cDNA sequenc
43	201.4	11.1	3371	22	AAH72766	Human cervical can
44	191.6	10.5	2484	22	AAF80536	Receptor #24 part1
45	187.6	10.3	1114	20	AAZ97687	Extended human sec

ALIGNMENTS

RESULT	1
AAZ90521	
ID	AAZ90521 standard; cDNA; 1824 BP.
XX	AAZ90521;
AC	AAZ90521;
XX	05-JUN-2000 (first entry)
DT	Human GPCR protein (HGPRP) encoding cDNA (clone ID 1258981).
XX	Human; G protein coupled protein receptor; HGPRP; cell proliferation; neurological; immune disorder; cytostatic; anti-arteriosclerotic;
DE	KW anti-atherosclerotic; hepatotropic; antiinflammatory; viricide; leukemia;
DE	KW immunomodulatory; anemia; asthma; gastrointestinal; anti-epileptic;
DE	KW anti-Alzheimer's; anti-Parkinsonian; gene therapy; ss.
XX	OS Homo sapiens.
XX	WO200015793-A2.
PN	23-MAR-2000.
XX	
PD	
XX	17-SEP-1999; 99WO-US20958.
PF	
XX	17-SEP-1998; 98US-0156513.
PR	
XX	(INCY-) INCYTE PHARM INC.
PA	
XX	Bandman O, Lal P, Tang YT, Corley NC, Guegler KJ, Gorgone GA;
PI	Baughn MR;
XX	WPI; 2000-271432/23.
DR	

; FILING DATE: 04-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Reiter, Stephen E.
; REGISTRATION NUMBER: 31,192
; REFERENCE/DOCKET NUMBER: FP41 9772
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-546-4737
; TELEFAX: 619-546-9392
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 906 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-367-264-2

Query Match 6.0%; Score 140; DB 3; Length 906;
Best Local Similarity 22.5%; Pred. No. 8.4e-06;
Matches 80; Conservative 50; Mismatches 140; Indels 86; Gaps 17;

QY	20	GAWAQGHVPPCCSQGLNPLYNLCDSRGANGIVLEAVAGAGIVTTFVLTILV--ASLPF	77
Db	568	GMW-----PNADLTGCEPIPVRYLEWSNIESIIAIAFSCGLILVTLFVTLFVLYRDPV	622
QY	78	VODTKRSLGLGTQVFFLGP-LGLFCLVFACVVKPDFSTCASRREFLEGVLFAICFSCLA	136
Db	623	VKSSREL-----YIILAGIFLGYVC-PFTLIARPTTSCYLRQLLVGLSSAMCYSA	675
QY	137	HVFALNFLARKNHG-----PR---GWVFTVALLTLVEVIINTEWLIITLVRGSG	184
Db	676	-VTKTNRIARILAGSKKICTRPREMSAWAQVIAISILISVOLT-----LVVTLI---	725
QY	185	EGGPQGNSSAGWAVASP-----CATANMDFVMAIYVMLLLGAFLGAW	228
Db	726	-----IMEPPMPILSYPSIKEVYLICNTSNLGVAPLGYNGLLIMSCITYAF	772
QY	229	-----PALCGRYKRWKKGHGVFLLTTATSVAIWVWIVMVTYGNKQHNSTWDDPTL	283
Db	773	KTRNVFANFNK-----YIAFTWYTTCTIIWLAFPYI-FGSNYKIITTCFAVLSVT	824
QY	284	LAANAWAFVLFYVI---PE---VSQVTKSSPEQSYOGD-MYPTRGVGYETILKEQK	332
Db	825	VALGCMFTPKMYIIIAKPERNVRSAPFTTSDVVRMHVGDGKLPGRSNTFLNIFRRKK	880

Search completed: September 16, 2002, 09:33:09
Job time: 168 sec


```

Query Match      6.1%; Score 142; DB 1; Length 1199;
Best Local Similarity 22.4%; Pred. No. 7.9e-06;
Matches 85; Conservative 54; Mismatches 153; Indels 88; Gaps 19;

Qy 20 GAWAGHQVPPCGSGGLNPLYNLCDRSGAWGIVLEAVAGAGIVTFTVLTILV--ASLPF 77
Db 568 GWPNAEU-----TGCEPIPVRYLEWSDIESIIAIAFSCGLIVTLFTVLIFVLYRTPV 622
Qy 78 VQDTKKRSLLTQVFFLLGT--LGFLCLVFAVCWRPDFSTCASRRFLFGVLFAIFCSCLAA 136
Db 623 VKSSREL-----YIILAGIFLGYVC-PFTLIARPTTSCYLQRLVGLSAMCYSA-- 675
Qy 137 HVFALNFARKNHG-----PR---GWVFTVALLTLTLEVIINTFEMIILITLVRSRG 184
Db 676 -VTKNRIARLAGKKIKCTKPKFNSAWAQVIIASILISVQLT-----LVVTLI----- 725
Qy 185 EGGQGHSSAGWAVASP-----CAIANMDFVMAIYYVMLLLGLGFLGAW 228
Db 726 -----IMEPPMPLTSPYSIKEVYLICNTSLNGVAVPYGVNGLLIMSCYYAF 772
Qy 229 -----PALCGRYKWRKHGVFLTLTATSVAIVWVMIYMTYGNKHQHSPTWDDPTLAI 283
Db 773 KTRNVPAFNEAK-----YIATMTYTCIIILWAFVPIY-FGSNYKIITTCFAVLSVT 824
Qy 284 LAANAWAFVLPYVI---PE---VSQVTKSSPEQSQGD-MYPTRGVGYETILKEOK-QQS 335
Db 825 VALCMTPTPKYIIIAKPERNVRSATFTSDVVRMHVGVGDKLPCRSNTFLNIFRRKKPGAG 884
Qy 936 MFVEN-KAFSMDPEPVAKR 354
Db 936 MFVEN-KAFSMDPEPVAKR 354

```

Query Match	6.18;	Score 142;	DB 1;	Length 1199;
Best Local Similarity	22.4%;	Pred. No. 7.9e-06;		
Matches	85; Conservative	54; Mismatches	153; Indels	88; Gaps
Qy	20	GAWAQGHVPPGCSQGLNPLYNLCDRSGAWGIIVLEAVAGAGITVFVTIILV--ASLPF	77	
Dd	568	GWPNNAEL-----TGCEPIPVRYLWSDIESIIAIAFSCLGILTFTLVLFVLVRTPTV	622	
Qy	78	VODTKRSLLGTQVFFLLGT-LIGVFLYFACVVPDFPSTCARRRPLFGVLFAICSCLA	136	
Dd	623	VKSSRELC---YIILAGIFLGYYC-PFTLTAKPTTSCYLQRULLVGLSAMCYSA-	675	
Qy	137	HVFALNFARKNHG-----PR--GVWTFVALLLTLEVINTEWLIITLVLRGSG	184	
Dd	676	-VTKNRTRARILAGSKKKICTRKPFMSANQAQVIATISILSVOLT-----LVVTLI----	725	
Qy	185	EGPGOGNSSAGWAVASP-----CAIANMDFVMALIYVMLLLGLGAFLGAW	228	
Dd	726	-----INEPMPIILSPSIREVDLICNTSLGNVVAPGYNGLLIMSCITYAF	772	

ATTORNEY/AGENT INFORMATION:
NAME: Blalock, Donna K.
REGISTRATION NUMBER: 38,082
REFERENCE/DOCKET NUMBER: X-8319B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317/277-1090
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1194 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-538-526-1

Query Match 6.38; Score 147; DB 4; Length 1194;
Best Local Similarity 20.68; Pred. No. 2.9e-06;
Matches 104; Conservative 66; Mismatches 192; Indels 142; Gaps 23;

QY 20 GAAQGHVPPCCSOGLNPLYNLCDRSGAWGIVLEAVAGAGIVTFTVLTIILV--ASLPPF 77
DB 568 GWW-----PNADLNGCEPIPVRYLEWSNIEPIIAIAFSCGLIVTLFVLFIYRDPV 622
QY 78 VQDTKKRSLGTVFFLLGT-LGLFCLVFCVWPKDFSTCASRRFLFGVLFAICFSCIAA 136
DB 623 VKSSSRELC-----YIILAGIFLGYVC-PFTLIAKPTTTSCYLQRLVLGLSSAMCYSAL-- 675
QY 137 HVFALNFIARKNHG-----PR---GWVFTVALLTLVEVIINTEWLIITLVGRSG 184
DB 676 -VTNTRNRIARLAGSKKIKCTKPRFMSAWAQVITIASILISVQLT-----LVVTLI---- 725
QY 185 EGGPGQNSSAGWAVASP-----CAIANMDFVMAIYVMLLLGLAFLGAW 228
DB 726 -----IWEPPMPILSYPSKEVYLICNTSLGVAAPVGVNGLIMSCYYAF 772
QY 229 -----PALCGRYKRWKHGVEVLLTATSAIWMVIMYTYGNKHNSPTWDDPTLAIA 283
DB 773 KTRNVNPFANEAK-----YIAFTMYTTCIIWLAFVPIY-FGSNYKIITTCFAVLSVT 824
QY 284 LAANAFVLFYVI---PE---VSQVTKSSPEQSQGD-MYPTRGVGVETILKEQ----- 331
DB 825 VALGCMETPKMYIIIAKPERNVRSAFTTSDVVRMHVGDGKLPKCRSNTFLNIFRRKKAGAG 884
QY 332 -----KGQSMF-----VENKAFSMDPEVAARPPVSPYSYNGOL 365
DB 885 NANSNGKSVSWSEPGGGOVPKGQHMWHRLSVHVKTNETACNQTAIVIKPLTKSYOG-SGKS 943
QY 366 LT-----SVYQPT-----MALMHKVPSEGAYDIILPRATANSQVMG 402
DB 944 LTFSDTSTKTLNVVEEEDAQPIRFPSPGSPSMVHRRVPSAATPPPLPHLTAEETPLF 1003
QY 403 SANSTLRAEDMYSAQSHQAATPPK 426
DB 1004 LAEPAL-PKGLPPLPQQOQPPQ 1026

RESULT 9
US-08-041-538-2
Sequence 2, Application US/08041538
Patent No. 5385831
GENERAL INFORMATION:
APPLICANT: Mulvihill, Eileen R
APPLICANT: Hagen, Frederick S
APPLICANT: Houamed, Khaled M
APPLICANT: Almers, Wolfhard
TITLE OF INVENTION: G PROTEIN COUPLED GLUTAMATE
NUMBER OF INVENTIONS: 15
CORRESPONDENCE ADDRESS:
ADDRESS: Townsend and Townsend
STREET: Steuart Street Tower, One Market Plaza
CITY: San Francisco
STATE: CA

COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/041,538
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/648,481
FILING DATE:
APPLICATION NUMBER: US 07/626,806
FILING DATE: 12-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Parmelee, Steven W
REGISTRATION NUMBER: 31,990
REFERENCE/DOCKET NUMBER: 13952-6-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-467-9600
TELEFAX: 206-623-6793
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1199 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-041-538-2

Query Match 6.18; Score 142; DB 1; Length 1199;
Best Local Similarity 22.48; Pred. No. 7.9e-06;
Matches 85; Conservative 54; Mismatches 153; Indels 88; Gaps 19;

QY 20 GAAQGHVPPCCSOGLNPLYNLCDRSGAWGIVLEAVAGAGIVTFTVLTIILV--ASLPPF 77
DB 568 GWWPNAEL-----TGCEPIPVRYLEWSNIEPIIAIAFSCGLIVTLFVLFIYRDPV 622
QY 78 VQDTKKRSLGTVFFLLGT-LGLFCLVFCVWPKDFSTCASRRFLFGVLFAICFSCIAA 136
DB 623 VKSSSRELC-----YIILAGIFLGYVC-PFTLIAKPTTTSCYLQRLVLGLSSAMCYSAL-- 675
QY 137 HVFALNFIARKNHG-----PR---GWVFTVALLTLVEVIINTEWLIITLVGRSG 184
DB 676 -VTNTRNRIARLAGSKKIKCTKPRFMSAWAQVITIASILISVQLT-----LVVTLI---- 725
QY 185 EGGPGQNSSAGWAVASP-----CAIANMDFVMAIYVMLLLGLAFLGAW 228
DB 726 -----IWEPPMPILSYPSKEVYLICNTSLGVAAPVGVNGLIMSCYYAF 772
QY 229 -----PALCGRYKRWKHGVEVLLTATSAIWMVIMYTYGNKHNSPTWDDPTLAIA 283
DB 773 KTRNVNPFANEAK-----YIAFTMYTTCIIWLAFVPIY-FGSNYKIITTCFAVLSVT 824
QY 284 LAANAFVLFYVI---PE---VSQVTKSSPEQSQGD-MYPTRGVGVETILKEQ-CQS 335
DB 825 VALGCMETPKMYIIIAKPERNVRSAFTTSDVVRMHVGDGKLPKCRSNTFLNIFRRKKPGAG 884
QY 336 MEVEN-KAFSMDPEVAARPP 354
DB 885 NANSNGKSVSWSEPGGROAP 904

RESULT 10.
US-08-463-642-2
Sequence 2, Application US/08463642
Patent No. 5721107
GENERAL INFORMATION:
APPLICANT: Mulvihill, Eileen R
APPLICANT: Hagen, Frederick S
APPLICANT: Houamed, Khaled M

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/072,574
; FILING DATE: 04-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Reiter, Stephen E.
; REGISTRATION NUMBER: 31,192
; REFERENCE/DOCKET NUMBER: FP41 9772
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-546-4737
; TELEFAX: 619-546-9392
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 879 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-486-270-6

Query Match      6.3%; Score 147; DB 1; Length 879;
Best Local Similarity 21.8%; Pred. No. 1.5e-06;
Matches 61; Conservative 55; Mismatches 106; Indels 58; Gaps 12;

Qy 48 AWGIVLEAVAGAGIVTTFTVLTIILVA--SLPFVQDTKKRSLLGTVQVFFLLGLTGL-FCLV 104
   ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 575 AWAIGPVTIACLGEMCTCMVTVTFIKHNNTPLVKASGRE-----LCYILLFGVGLSYCMT 629

Qy 105 FACVVKPDFSTCASRFLFGVLFALFSCSLAAHVFAFNFLAR-----KNHGPRGWI--- 156
   ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 630 FFFIAKSPVICALRRLGSSFAICYALLTKT---NCIARIFDGVKNGAQRKPFISPS 686

Qy 157 --FTVALLTLVEVLIINTEWLIITLVRSGBGGPGQNSAGWAVAS-----PCAIANM 207
   ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 687 SQVFICGLILVQIVMVSVWLIL-----EAPGTRYTLAEKRETVILKCNVKDS 735

Qy 208 DFVMALIVYMLLLGAFGLGAWPALCGRYK-RWRK-----HGVEVLLTTATSVAIWVWI 260
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 736 SMLISLTVDVILVI-----LCTVYAFKTRKCPENFNEAKFIGFTMTYTTCIWLAF 786

Qy 261 VMYTYGNKQHSPTWDDPTLAIALAANAWAFVLFYVIPEV 300
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 787 PIFYTSSDIRVQT---TTMCISVSLSGFVVLGCLFAPKV 823
```

```
RESULT 7
US-08-267-264-6
; Sequence 6, Application US/08367264
; Patent No. 6001581
; GENERAL INFORMATION:
; APPLICANT: Daggett, Lorrie
; APPLICANT: Ellis, Steven B.
; APPLICANT: Liaw, Chen
; APPLICANT: Pontsler, Aaron
; APPLICANT: Johnson, Edwin C.
; APPLICANT: Hess, Stephen D.
; TITLE OF INVENTION: HUMAN METABOTROPIC GLUTAMATE RECEPTORS,
; NUCLEIC ACIDS ENCODING SAME AND USES THEREOF
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
; STREET: 444 South Flower Street, Suite 2000
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/367,264
; FILING DATE: 02-JUN-1994
```

```
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/072,574
; FILING DATE: 04-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Reiter, Stephen E.
; REGISTRATION NUMBER: 31,192
; REFERENCE/DOCKET NUMBER: FP41 9772
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-546-4737
; TELEFAX: 619-546-9392
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 879 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-367-264-6

Query Match      6.3%; Score 147; DB 3; Length 879;
Best Local Similarity 21.8%; Pred. No. 1.5e-06;
Matches 61; Conservative 55; Mismatches 106; Indels 58; Gaps 12;

Qy 48 AWGIVLEAVAGAGIVTTFTVLTIILVA--SLPFVQDTKKRSLLGTVQVFFLLGLTGL-FCLV 104
   ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 575 AWAIGPVTIACLGEMCTCMVTVTFIKHNNTPLVKASGRE-----LCYILLFGVGLSYCMT 629

Qy 105 FACVVKPDFSTCASRFLFGVLFALFSCSLAAHVFAFNFLAR-----KNHGPRGWI--- 156
   ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 630 FFFIAKSPVICALRRLGSSFAICYALLTKT---NCIARIFDGVKNGAQRKPFISPS 686

Qy 157 --FTVALLTLVEVLIINTEWLIITLVRSGBGGPGQNSAGWAVAS-----PCAIANM 207
   ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 687 SQVFICGLILVQIVMVSVWLIL-----EAPGTRYTLAEKRETVILKCNVKDS 735

Qy 208 DFVMALIVYMLLLGAFGLGAWPALCGRYK-RWRK-----HGVEVLLTTATSVAIWVWI 260
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 736 SMLISLTVDVILVI-----LCTVYAFKTRKCPENFNEAKFIGFTMTYTTCIWLAF 786

Qy 261 VMYTYGNKQHSPTWDDPTLAIALAANAWAFVLFYVIPEV 300
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 787 PIFYTSSDIRVQT---TTMCISVSLSGFVVLGCLFAPKV 823

RESULT 8
US-08-538-526-1
; Sequence 1, Application US/08538526
; Patent No. 6303751
; GENERAL INFORMATION:
; APPLICANT: Burnett, J. Paul
; APPLICANT: Mayne, Nancy G.
; APPLICANT: Sharp, Robert L.
; APPLICANT: Snyder, Yvonne M.
; TITLE OF INVENTION: Human Metabotropic Glutamate Receptor
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patent Division/DKB
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/538,526
; FILING DATE: October 3, 1995
; CLASSIFICATION: 530
```

```
; APPLICATION NUMBER: 08/337,797
; FILING DATE: No. 6103475ember 14, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Gaylo, Paul J.
; REGISTRATION NUMBER: 36,808
; REFERENCE/DOCKET NUMBER: X-9431
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317) 276-0756
; TELEFAX: (317) 276-3861
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 872 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-258-523-2
```

```
Query Match 6.6%; Score 153; DB 3; Length 872;
Best Local Similarity 22.2%; Pred. No. 3.3e-07;
Matches 75; Conservative 50; Mismatches 143; Indels 70; Gaps 14;

Qy 6 ALVNCGLP--LFUFGAWAGHVPPGCSQGLNPLYYNLCDSRGAWGIVLEAVAGAGIVT 63
Db 549 SLTGCFFELPQEIYRWGDWAVGPVTIAC-----LGA-LAT 582

Qy 64 TFVLTITLVASLPVQDTKRRSLGTQVFFLLGTLGL-FCLVFACVVKVPDEFSTCASRRFL 122
Db 583 LFVLGVFVRINATPVVWASGREL----CYLLGGVFLCYCMTFIFIAKPSGTGVCALRRLG 638

Qy 123 FGVLFACFCFLAAHVFALFNLRKHGPRGW-----IFVALLLTIV--EVIINT 172
Db 639 VGTAFSVCYSAALLTKT---NRIAFGAREGAORPRFISPASOVAICLALISQQLIVV 695

Qy 173 EWLITLVRSGEPPGCGNSAGWAVASPCAIANMDFMALIYVMLLLGAFLGAWPALC 232
Db 696 AWWVEAPGKGKETAPERRE----VTLRCNHRDASMLGSLAYNVLLI-----ALC 742

Qy 233 GRYK-RWRK-----HGVFVLLTATSVAIWVWVMTYGNKOHNSPTWDDPTLATALA 285
Db 743 TLVAFKTRKCPENFNEAKFGFTMTTCIILWLAFLPIFYVTSSDYRVQT---TTMCSVSVS 799

Qy 286 ANAWAFVLFVIVPEVSQVTKSSPEQSQGDMYPTRGVG 323
Db 800 LSG-SVVLGCLFAPKHLIILFQPKNVVSHRAPTSREG 836
```

```
RESULT 5
US-08-072-574-6
; Sequence 6, Application US/08072574
; Patent No. 5521297
; GENERAL INFORMATION:
; APPLICANT: Daggett, Lorrle
; APPLICANT: Ellis, Steven B.
; APPLICANT: Liaw, Chen
; APPLICANT: Pontsler, Aaron
; TITLE OF INVENTION: HUMAN METABOTROPIC GLUTAMATE RECEPTORS,
; NUCLEIC ACIDS ENCODING SAME AND USES THEREOF
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
; STREET: 444 South Flower Street, Suite 2000
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 00719
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/072,574
```

```
; FILING DATE: 19930604
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Reiter, Stephen E.
; REGISTRATION NUMBER: 31,192
; REFERENCE/DOCKET NUMBER: P41 9383
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 213-622-7700
; TELEFAX: 213-489-4210
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 879 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-072-574-6

Query Match 6.3%; Score 147; DB 1; Length 879;
Best Local Similarity 21.8%; Pred. No. 1.5e-06;
Matches 61; Conservative 55; Mismatches 106; Indels 58; Gaps 12;

Qy 48 AWGIVLEAVAGAGIVTTFVLTIILVA--SLPFVQDTKRRSLLGTVQVFFLLGTLGL-FCILV 104
Db 575 AWAIGPVTTIACLGPMCTCMVVTVFVKHNTPLVKASGRE-----LCYILLFGVGLSYCMT 629

Qy 105 FACVVKPDFSTCASRRFLFGVLFPAICFSCLAHVFALNFLAR-----KNHGPRGWV---- 156
Db 630 FFETAKPSVICALURRLGLSGSSFAICYSAALLTKT---NCTARIFDGKVGKGAQRKPFISPS 686

Qy 157 --FTVALLLTLVEVIIINTEWLIIITLVRSGBGGPOGNSAGWAVAS-----PCAIANM 207
Db 687 SQVFICGLILVQIVWVSVMLIL-----EAGTRRYTLAEKRETVILKCNKVD 735

Qy 208 DFVMAIYVMLLLGAFLGAWPALCGRYK-RWRK-----HGVFVLLTATSVAIWVWVI 260
Db 736 SMLISLYDVILVI-----LCTVYAPKTRKCPENFNEAKFGFTMTTCIILWLAFL 786

Qy 261 VMTYGNKOHNSPTWDDPTTALIAAANAWAFVLFYVPEV 300
Db 787 PIIFYTSSDYRVQT---TTMCISVSLSGFVVVLGCLFAPKV 823
```

```
RESULT 6
US-08-486-270-6
; Sequence 6, Application US/08486270
; Patent No. 5807689
; GENERAL INFORMATION:
; APPLICANT: Daggett, Lorrle
; APPLICANT: Ellis, Steven B.
; APPLICANT: Liaw, Chen
; APPLICANT: Pontsler, Aaron
; APPLICANT: Johnson, Edwin C.
; APPLICANT: Hess, Stephen D.
; TITLE OF INVENTION: HUMAN METABOTROPIC GLUTAMATE RECEPTORS,
; NUCLEIC ACIDS ENCODING SAME AND USES THEREOF
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
; STREET: 444 South Flower Street, Suite 2000
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,270
; FILING DATE: 02-JUN-1994
; CLASSIFICATION: 435
```


Db 271 -----DAFCKPQLVKKSYGVENRAYSQEE-----ITQGFETGDTLYAPYS-THFQ 315

QY 377 LMHKVPSEGAYDILPRATA 396

Db 316 LQNOPPOK-----EFSIPRAHA 332

RESULT 2

US-09-188-930-123

Sequence 123, Application US/09188930A

Patent No. 6150502

GENERAL INFORMATION:

APPLICANT: Watson, James D.

APPLICANT: Strachan, Lorna

APPLICANT: Sleeman, Matthew

APPLICANT: Onrust, Rene

APPLICANT: Murison, James Greg

TITLE OF INVENTION: Compositions Isolated From Skin Cells

TITLE OF INVENTION: and Methods For Their Use

FILE REFERENCE: 11000.1011c1

CURRENT APPLICATION NUMBER: US/09/188,930A

CURRENT FILING DATE: 1998-11-09

NUMBER OF SEQ ID NOS: 348

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 123

LENGTH: 68

TYPE: PRT

ORGANISM: Human

US-09-188-930-123

Query Match

Best Local Similarity 7.3%; Score 169; DB 4; Length 68;

Matches 38; Conservative 11; Mismatches 17; Indels 2;

QY 66 VLIT-ILVASLPFVDTKKSLGTQVFLGLTGLFCLVFCVVKPFDFTCSRRFLFG 124

Db 1 MLTLPILVCK--VQDSNRNRKMLPTQFLFLGLVGLFGLTFAFIIGLDGSGTGTREPLFG 57

QY 125 VLFAICFSCSL 134

Db 58 ILFSICFSCSL 67

RESULT 3

US-08-337-797A-2

Sequence 2, Application US/08337797A

Patent No. 6017697

GENERAL INFORMATION:

APPLICANT: Burnett, J. P.

APPLICANT: Mayne, Nancy G.

APPLICANT: Sharp, Robert L.

APPLICANT: Snyder, Yvonne M.

TITLE OF INVENTION: EXCITATORY AMINO ACID RECEPTOR PROTEIN

TITLE OF INVENTION: AND RELATED NUCLEIC ACID COMPOUNDS

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESSEE: Eli Lilly and Company

STREET: Lilly Corporate Center

CITY: Indianapolis

STATE: Indiana

COUNTRY: United States of America

ZIP: 46285

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/337,797A

FILING DATE: No. 6017697ember 14, 1994

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Gaylo, Paul J.

REGISTRATION NUMBER: 36,808

REFERENCE/DOCKET NUMBER: X-9431

TELECOMMUNICATION INFORMATION:

TELEPHONE: (317) 276-0756

TELEFAX: (317) 276-3861

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 872 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-337-797A-2

Query Match 6.6%; Score 153; DB 3; Length 872;

Best Local Similarity 22.2%; Pred. No. 3.3e-07;

Matches 75; Conservative 50; Mismatches 143; Indels 70; Gaps 14;

QY 6 ALVMCLGLP--LFLFPGAWAQHVPPGCSQGLNPLYYNLCDRSGAWGIVLEAVAGAGIVT 63

Db 549 SLTGCFELPQEIYRWGDWAVGPVTIAC-----LGA-LAT 582

QY 64 TFVLTILVASLPFVDTKKRSLLGTQVFPFLGLTGL-FCLVFCVVKPFDFTCSRRRL 122

Db 583 LEVLGVFVRHNATPVVKASGREL---CYLLGGVFLCYCMTEFIAKPSGTGVCALRRLG 638

QY 123 FGVLFATCFSCSLAAHVFAFNFLARKNHGPRGW-----IFTVALLLTV--EVIINT 172

Db 639 VGTAFSVYCSALLTKT---NRIARIFGGAREGAQRFFISPASQVAICLALISGQLLIV 695

QY 173 EWLITLVRGSGEGPGQNSAGWAVASPCAIAIMDFVMAIYVMLLLGLGAFLGAWPALC 232

Db 696 ARLVVEAPGPGKETAPERRE---VVTLCRHRDASMLGSLAYNVLLI-----ALC 742

QY 233 GRVK-RWRK-----HGVFVLLTATSAIVAIWVWVWYVWYVWYVWYVWYVWYVWY 285

Db 743 TLYAFKTRKCPENFENAKFTGFTMYTTCIIWLAFLPIFYVTSSDYRVQT---TTMCVSVS 799

QY 286 ANAWAFVLFVYVPEVSQVTKSSPEQSYQGDMYPTRGV 323

Db 800 LSG-SVVLGCLFAPKLHIILFQKNNVSHRAPTSRFG 836

RESULT 4

US-09-258-523-2

Sequence 2, Application US/09258523

Patent No. 6103475

GENERAL INFORMATION:

APPLICANT: Burnett, J. P.

APPLICANT: Mayne, Nancy G.

APPLICANT: Sharp, Robert L.

APPLICANT: Snyder, Yvonne M.

TITLE OF INVENTION: EXCITATORY AMINO ACID RECEPTOR PROTEIN

TITLE OF INVENTION: AND RELATED NUCLEIC ACID COMPOUNDS

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESSEE: Eli Lilly and Company

STREET: Lilly Corporate Center

CITY: Indianapolis

STATE: Indiana

COUNTRY: United States of America

ZIP: 46285

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/258,523

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2002, 09:30:21 ; Search time 15.84 Seconds
(without alignments)
680.031 Million cell updates/sec

Title: US-09-895-686-1
Perfect score: 2326
Sequence: 1 MAIHKALVCMGLGLFLFPG.....ATPKDGKNSQVFRNPYWD 441

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues
Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	533	22.9	347	4 US-09-188-930-326	Sequence 326, App
2	169	7.3	68	4 US-09-188-930-123	Sequence 123, App
3	153	6.6	872	3 US-08-337-797A-2	Sequence 2, Appli
4	153	6.6	872	3 US-09-258-523-2	Sequence 2, Appli
5	147	6.3	879	1 US-08-072-574-6	Sequence 6, Appli
6	147	6.3	879	1 US-08-486-270-6	Sequence 6, Appli
7	147	6.3	879	3 US-08-367-264-6	Sequence 6, Appli
8	147	6.3	1194	4 US-08-538-526-1	Sequence 1, Appli
9	142	6.1	1199	1 US-08-041-538-2	Sequence 2, Appli
10	142	6.1	1199	1 US-08-463-642-2	Sequence 2, Appli
11	142	6.1	1199	1 US-08-455-602-2	Sequence 2, Appli
12	142	6.1	1199	2 US-08-461-157-2	Sequence 2, Appli
13	142	6.1	1199	5 PCT-US91-09422-2	Sequence 2, Appli
14	140	6.0	906	1 US-08-486-270-2	Sequence 2, Appli
15	140	6.0	906	3 US-08-367-264-2	Sequence 2, Appli
16	136.5	5.9	1219	2 US-08-687-289A-6	Sequence 6, Appli
17	136	5.8	906	5 PCT-US91-09422-17	Sequence 17, Appli
18	131.5	5.7	1056	2 US-08-687-289A-7	Sequence 7, Appli
19	130.5	5.6	1056	2 US-08-687-289A-8	Sequence 8, Appli
20	125	5.4	908	4 US-08-855-146-2	Sequence 2, Appli
21	123	5.3	1058	2 US-08-687-289A-5	Sequence 5, Appli
22	122	5.2	908	3 US-08-823-110-1	Sequence 1, Appli
23	122	5.2	908	3 US-08-604-298-1	Sequence 1, Appli
24	118.5	5.1	877	3 US-09-126-280-2	Sequence 2, Appli
25	118.5	5.1	1079	1 US-08-485-588-8	Sequence 8, Appli
26	118.5	5.1	1079	1 US-08-484-565-8	Sequence 8, Appli
27	118.5	5.1	1079	2 US-08-480-751-8	Sequence 8, Appli

28	118.5	5.1	1079	2 US-08-943-986-8	Sequence 8, Appli
29	118.5	5.1	1079	3 US-08-353-784-8	Sequence 8, Appli
30	118.5	5.1	1079	3 US-08-484-719B-8	Sequence 8, Appli
31	118.5	5.1	1079	4 US-08-484-159-8	Sequence 5, Appli
32	118	5.1	1085	1 US-08-485-588-5	Sequence 5, Appli
33	118	5.1	1085	1 US-08-484-565-5	Sequence 5, Appli
34	118	5.1	1085	2 US-08-480-751-5	Sequence 5, Appli
35	118	5.1	1085	2 US-08-943-986-5	Sequence 5, Appli
36	118	5.1	1085	3 US-08-353-784-5	Sequence 5, Appli
37	118	5.1	1085	4 US-08-484-719B-5	Sequence 5, Appli
38	118	5.1	1085	4 US-08-484-159-5	Sequence 2, Appli
39	117.5	5.1	1027	4 US-09-162-021B-2	Sequence 2, Appli
40	117.5	5.1	1078	1 US-08-485-588-7	Sequence 7, Appli
41	117.5	5.1	1078	1 US-08-484-565-7	Sequence 7, Appli
42	117.5	5.1	1078	2 US-08-480-751-7	Sequence 7, Appli
43	117.5	5.1	1078	2 US-08-943-986-7	Sequence 7, Appli
44	117.5	5.1	1078	3 US-08-353-784-7	Sequence 7, Appli
45	117.5	5.1	1078	3 US-08-484-719B-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-09-188-930-326
; Sequence 326, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 326
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Human
US-09-188-930-326

Query Match	22.9%	Score 533;	DB 4;	Length 347;
Best Local Similarity	36.1%	Pred. No. 4.7e-48;		
Matches 137;	Conservative 61;	Mismatches 130;	Indels 52;	Gaps 12;
QY	21	AWAQHVPPGCSQGLNPLYYNLCDSRGAMGIVLEAVAGAGIVTT--FVLTI-ILVASLPF	77	
Db	1	AWSR-----PRYRLCDKAEAGIVLETATAGVTVSVAFMLTLPILVCK---	45	
QY	78	VQDKRSLLGTQVFLGLTGLFCLVFACVVKPDPSTCASRRFLGVLFAICFSLAAH	137	
Db	46	VQDSNRKMLPTQFLGLLGLVGLGIFGLTFAIIGLDGSGTGRFFLGLFGLFICFSCLLAH	105	
QY	138	VFALNFLARKNHGPRGWITFTVALLTLVEIINTWLIITLVRGSGEGGPGQNSAGWA	197	
Db	106	AVSLTKLVGRKPLSLVILGLAVGSLVQDVIAIEYIVLTMR-----TNVWFSE	157	
QY	198	VASPCAANMDFVMALYYVMLLLGALGAWPALCGRYKRWKRGHGVFVLLTTATSVAIWV	257	
Db	158	LSAP--RRNEDEVLLTYVFLMALTFMSSFTFCGSGFTGWRKHGAHIVLTMLLSIAIWV	215	
QY	258	VWIVMTYGNKQUNSTWDDPTLAIANAANAFVLEFYIPEVSQVTKSPESQSGODMY	317	
Db	216	AWTLLMLPDFDR---WDDTILSSALAANGWVFLLAYVSPFWLLTKORNPMDYDVE--	270	
QY	318	PTRGVGYETILKEQGSME-VENKAFSMDPEVAAKRVPSPYSGYNGQLTTSYVQPTEMA	376	

Search completed: September 16, 2002, 09:32:47
Job time: 206 sec

RESULT 15

thio-disulfide interchange protein, probable [imported] - Caulobacter crescentus

Query Match

12 GLPLFLFPG-AWAOGHVPPGCSOGI.NPLYYNI.CDRSGAWGTVI.---EAVAGA----- 59

Best Local Similarity

Matches 105; Conservative 57; Mismatches

12 G L P I E I F P G - A W A O G H V P P G C S O G I N P I Y Y N I C D R S G A W G T V I . - - - - E A V A G A - - - - - 59

[illegible]

223 GLTLKPGYDFVGGSTPPAELVGV-----LATKAGAWEVTTAGEALAGASGLGPPPA 276

60 -----GIVTTEVI TTTI VASI BEVODTKKRSI I CTOVEFI I CTICI ECI VEACWIKDDDES 114

[illegible]

277 EAAPGALA-----GGLAGALLFAFLG--GLILNLMPGVF-PVLS 313

115 TCASB-----DET ECVI EA TCESCI AADVEAY NET A DYNUCDDCW-----15A

[illegible]

314 MKAASLAGHAHEAPKARLQGLAFLGVV--ATFLALAGALLAV---RAGGAAVGWGEQL 367

[illegible]

133 ---VIFVALLLEVLVEVINIENWLLIILVRG-----SGEGGPQG---NSSAGWAVASPC 202

368 QSPLVIAGLALLMLLVALNMSGIFEIGTSVQGVGAGASAKGVSGAFFTGALAVVVAAPC 427

[illegible]

203 ATANMDFVM--ALTYVMLLLGAF LG-----AWPALCGRYKRWKRHHGVFV----- 245

428 TAPFMAGALGYALTQPPVIALGVFLALALGFAAPEFVAVTFIPGALKLLPRPGDWMEVLKK 487

246 -LLTATSAIWWIV-----MYTYGNKQHNSPTWDDPTLAI 282

488 GLAFPMYGAALWLWVFAQQAGPIALGQLLVAGVLAAFGAWLYGLAQARRAVGKGS--AV 545

283 ALAANAWAFVLFYVIPEVSQVTKSSPEQSYQCDMPYTRGVGYETILKEQGQSMFVENKA 342

546 SMILGLLAVVGALALAAASALSASAKPPVAAAEEASTPSGPGLTAEA-WSPEKVOALOAEGRP 604

343 FSMD-----EPVAAKRP--VSPYSGYNGQLLTSVYQPTMALMHKV----- 381

A;Cross-references: GB:L08132; NID:g290215; PIDN:AAA72332.1; PID:g290216
C;Genetics:
A;Gene: FlyBase:DvIr/boss
A;Cross-references: FlyBase:FBgn0013105
A;Introns: 14/3; 175/3; 324/3; 808/3
C;Keywords: transmembrane protein

Query Match 6.3%; Score 146; DB 2; Length 893;
Best Local Similarity 20.8%; Pred. No. 0.0006;
Matches 71; Conservative 58; Mismatches 136; Indels 76; Gaps 15;

QY 34 GLNPYYNLCDRSCA--WGIVLEAVAGAGIVTTPVLTILLVASLPFFQDTKKRSLLG--- 88
 | | :
DB 503 GENRRYPFLFDGESVMWRKLDTWATTAALGIATLAILVFI---VVRISLGDNVF 559
 | | :
QY 89 -----TQVFFLLGTGLGFC--LVFAC-----VVKPDFST----CASRRRLFGVLV 127
 | :
DB 560 EGNPVTSLILLLSILVFCSPVPFSMEYVGQRNSHTTFEDVHTLNTLCGRVRFIMTLVY 619
 | :
QY 128 AICFSCIAAHVFAL-----NFLARNHGRGWIFTVALLLTVEVIINTEWLIITLV 180
 | :
DB 620 CFVESLLLCRAVMLASIGSEGGFLSHVN---GYIQAIIICVLSVFVGMSVQLLVV--M 673
 | :
QY 181 RSGEGGPGQNSSAG-WAVASPCAIANDFVVALIYVMLLLGAFLGAWPALCGRYRW 239
 | :
DB 674 HLASESVSCENIYGRW-----LWGLLAYDFDLCLSLVSLVPFYIRSQRNY 720
 | :
QY 240 KHGVFVLLTTATSAIVWVVMTYTYGNKHQSPTDDPTLAIALAANAFAVLFYVIP- 298
 | :
DB 721 E-GILIVIGAVLIILINSWTALSNGDE-----WRDAIPLGMQASGNVAIVGILIPR 773
 | :
QY 299 -----EVSQVTKSSPEQS-----YQGDMPYTRGVGYETI 327
 | :
DB 774 TFLIVRGIERSDIAQLPSLTSLAFAQNNQYSSESQSYVECV 814
 | :

RESULT 9
A41939
G protein-coupled glutamate receptor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
C;Accession: A41939; S15362
R;Houamed, K.M.; Kuiper, J.L.; Gilbert, T.L.; Haldeman, B.A.; O'Hara, P.J.; Mulvihill
Science 252, 1318-1321, 1991
A;Title: Cloning, expression, and gene structure of a G protein-coupled glutamate rec
A;Reference number: A41939; MUID:92022526
A;Accession: A41939
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-1199 <HOU>
A;Cross-references: GB:M61099; NID:g397806; PIDN:AAA19497.1; PID:g204460
A;Experimental source: cerebellum
A;Note: sequence extracted from NCBI backbone (NCBIP:60785)
R;Masu, M.; Tanabe, Y.; Tsuchida, K.; Shigemoto, R.; Nakanishi, S.
Nature 349, 760-765, 1991
A;Title: Sequence and expression of a metabotropic glutamate receptor.
A;Reference number: S15362; MUID:91156047
A;Accession: S15362
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1199 <MAS>
A;Cross-references: EMBL:X57569; NID:g56646; PIDN:CAA40799.1; PID:g56647
C;Keywords: G protein-coupled receptor; transmembrane protein

Query Match 6.1%; Score 142; DB 2; Length 1199;
Best Local Similarity 22.4%; Pred. No. 0.0018;
Matches 85; Conservative 54; Mismatches 153; Indels 88; Gaps 19;

QY 20 GAAAGCHVPPGCQGLNPLYYNLCDRSGAMGIVLEAVAGAGIVTFTVLTILV--ASLPF 77
 | | :
DB 568 GWPNNAEL-----TCCEPIPVRYLESWDIESITAFACSLGILFLFTFLVLYRDPPV 622
 | | :

Best Local Similarity 19.8%; Pred. No. 5.1e-05;
Matches 69; Conservative 63; Mismatches 130; Indels 86; Gaps 14;

QY 32 SQGLNPLYNLCDRSGA--WGIVLEAVAGAGIVTTFTVLITILVASLPFVQDTKKRSLLG- 88
Db 504 SAGDNNRRYPFLDGESVMFWRIKMDTWATGLTAAILGLIATLAILVFI---VVRISLGD 560

QY 89 -----TQVFLLGLTGLFCLVF-----ACVVKPDFST-----CASRRFLFGV 125
Db 561 VFEGNPTTSILLLSLILVFCSPVPYSIEYVGEQRNSHTVFEDAQTLNTLCVAVRFIMTL 620

QY 126 LFAIFCSCLAAHFAL-----NFLARKNHGPRGWIFTVALLLLTIVEIINTEWLIIT 178
Db 621 VYCFVFSLLLCRAVMLASIGSEGGFLSHVN---GIYQAVICAFSVVAQVGMVSQVLLVVM 676

QY 179 LVKSGEGGPGQNSSAGWAVASPCATANMDF-----VMALIYVMLLLLLGAFILGAWPALC 232
Db 677 HV-----ASETVCSENIIYGRWLWGLLAYDFALLCCVGALI---PSIY 716

QY 233 GRYKRRKHGCVFVLLTTATSVAIWVWVIMVYTYGNKQHSPTWDDPTLATAAANAWAFV 292
Db 717 RSNQRTRE-GILIVIGSVLIMVIAWIALSLFGDE-----WRDAIPLGLQASGAWVL 769

QY 293 LFYVIP-----EVSQVTKSPRQS-----YQGDWYPTRGVGYETI 327
Db 770 VGLIPRTFLIVRGIERSDIAQALPSLTSIAFAQNNQSSQSVYECV 817

RESULT 5
S26740
gene boss protein - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 16-Feb-1997
C:Accession: S26740
R:Kramer, H.
submitted to the EMBL Data Library, September 1990
A:Reference number: S26740
A:Accession: S26740
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-896 <KRA>
A:Cross-references: EMBL:X55887
C:Genetics:
A:Gene: FlyBase:boss
A:Cross-references: FlyBase:FBgn0000206
A:Introns: 18/1; 175/1; 325/1; 811/3

Query Match 6.7%; Score 157; DB 2; Length 896;
Best Local Similarity 20.5%; Pred. No. 7.4e-05;
Matches 87; Conservative 65; Mismatches 166; Indels 106; Gaps 19;

QY 32 SQGLNPLYNLCDRSGA--WGIVLEAVAGAGIVTTFTVLITILVASLPFVQDTKKRSLLG- 88
Db 504 SAGDNNRRYPFLDGESVMFWRIKMDTWATGLTAAILGLIATLAILVFI---VVRISLGD 560

QY 89 -----TQVFLLGLTGLFCLVF-----ACVVKPDFST-----CASRRFLFGV 125
Db 561 VFEGNPTTSILLLSLILVFCSPVPYSIEYVGEQRNSHTVFEDAQTLNTLCVAVRFIMTL 620

QY 126 LFAIFCSCLAAHFAL-----NFLARKNHGPRGWIFTVALLLLTIVEIINTEWLIIT 178
Db 621 VYCFVFSLLLCRAVMLASIGSEGGFLSHVN---GIYQAVICAFSVVAQVGMVSQVLLVVM 676

QY 179 LVKSGEGGPGQNSSAGWAVASPCATANMDF-----VMALIYVMLLLLLGAFILGAWPALC 232
Db 677 HV-----ASETVCSENIIYGRWLWGLLAYDFALLCCVGALI---PSIY 716

QY 233 GRYKRRKHGCVFVLLTTATSVAIWVWVIMVYTYGNKQHSPTWDDPTLATAAANAWAFV 292
Db 717 RSNQRTRE-GILIVIGSVLIMVIAWIALSLFGDE-----WRDAIPLGLQASGAWVL 769

QY 293 LFYVIP-----EVSQVTKS-----SPEQSYQGDWYPTRGV---YETILKEQG 333

Db 770 VGLIPRTFLIVRGIERSDIAQALPSLTSIAFAPEQSI-----LLGTGSVYECVNPAMRH 823
QY 334 QSMFVENKAESDEPVAAKRPVSPYSGYNQQLTTSVQPTMELMHKV-----PSEGAYDI 389
Db 824 CSODEVNHQSPSEIPTPLRGGGP---RRQOFFANLRQPMPTLIHNVHLDPSPVPPD 879

QY 390 ILPR 393
Db 880 RCPR 883

RESULT 6
JH0561
metabotropic glutamate receptor 2 precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 13-Sep-1998
C:Accession: JH0561
R:Tanabe, Y.; Masu, M.; Ishii, T.; Shigemoto, R.; Nakanishi, S.
Neuron 8, 169-179, 1992
A:Title: A family of metabotropic glutamate receptors.
A:Reference number: JH0561; MUID:92110002
A:Accession: JH0561
A:Molecule type: mRNA
A:Residues: 1-872 <TAN>
A:Experimental source: brain
C:Comment: This protein is coupled to a G protein and evokes a variety of functions b
C:Superfamily: metabotropic glutamate receptor 4
C:Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane p
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-872/Product: metabotropic glutamate receptor 2 #status predicted <MET>
F:568-590/Domain: transmembrane #status predicted <TRI>
F:605-625/Domain: transmembrane #status predicted <TII>
F:637-655/Domain: transmembrane #status predicted <III>
F:680-700/Domain: transmembrane #status predicted <TIV>
F:726-747/Domain: transmembrane #status predicted <TRV>
F:761-782/Domain: transmembrane #status predicted <TVI>
F:795-819/Domain: transmembrane #status predicted <VII>
F:203,286,338,402,547/Binding site: carboxylate (Asn) (covalent) #status predicted
F:601,675,827,837,843/Binding site: phosphate (Ser) (covalent) #status predicted
F:832/Binding site: phosphate (Thr) (covalent) #status predicted

Query Match 6.5%; Score 152; DB 2; Length 872;
Best Local Similarity 21.9%; Pred. No. 0.00019;
Matches 74; Conservative 51; Mismatches 143; Indels 70; Gaps 14;

QY 6 ALVMCLGLP--LFLFPGAWAQGHVPPGCSQGLNPLNYNLCDRSGAMGIVLEAVAGAGIVT 63
Db 549 SLTGCFELPQEIYRWGDAMAVGPVTIAC-----LGA-LAT 582

QY 64 TFLVTILVASLPFVQDTKKRSLLGTVQVFLGLTGL-FCLVFACVVKPDPFSTCASRRFL 122
Db 583 LFLVGFVRHNATPVVKASGREL---CYILGGVFLCYCMTEVFIKAPSTAVCTLRLRG 638

QY 123 FGVLEAIFCSCLAAHFALFNFLARKNHGPRGWV-----IFTVALLTLV---EVLINT 172
Db 639 LGTAFSVCSYALLTKT---NRIARIFGAREGAGRPRFTSPASQVAICLALISGQLLIVA 695

QY 173 EWLIITLVKSGEGGPGQNSSAGWAVASPCAIANMDFVMALIYVMLLLLLGAFILGAWPALC 232
Db 696 AWLVEAPGTGKETAPERRE---VVTLCNRHDSMLGSLAYNVLLI-----ALC 742

QY 233 GRYK-RWRK-----HGVFVLLTATSVAIWVWVIMVYTYGNKQHSPTWDDPTLATA 285
Db 743 TLVAFTRKCPENFNEAKFEGFTMYTTCIIWLAFLPIFYVTSDDYRVQT---TTCVSVS 799

QY 286 ANAWAFVLFVPIEVSQVTKSSPEQSYQGDWYPTRGV 323
Db 800 LSG-SVVLGCLFAPKLHILFQPKNVVSHRAPTSRFG 836

RESULT 7

[illegible]

C;Keywords: differentiation; G protein-coupled receptor; receptor; transmembrane prot
F;1-24/Domain: signal sequence glutamate receptor <Sig>
F;25-879/Product: metabotropic glutamate receptor subtype 3 #status predicted <TM1>
F;577-599/Domain: transmembrane #status predicted <TM1>
F;614-634/Domain: transmembrane #status predicted <TM2>
F;646-664/Domain: transmembrane #status predicted <TM3>
F;689-709/Domain: transmembrane #status predicted <TM4>
F;735-756/Domain: transmembrane #status predicted <TM5>
F;770-791/Domain: transmembrane #status predicted <TM6>
F;804-828/Domain: transmembrane #status predicted <TM7>

Query Match 6.9%; Score 160.5; DB 2; Length 879;
Best Local Similarity 22.5%; Pred. No. 3.7e-05;
Matches 69; Conservative 56; Mismatches 123; Indels 59; Gaps 14;

QY 19 PGAAOQHVPVPGSGQLNPLYNLCDR-----SGAWGIVLEAVAGAGIVTFTVLILVA- 73
Db 551 PGOWPTADL-SGC-----YNLPEDYTRWEDNAWGVTIACLGFMCTCIVIVTFKH 601

QY 74 -SLFFVQDTKRSLLTQTQVFLL--GTGLGLFCLVACVVVKPDFSTCASRRFLFGVLFAIC 130
Db 602 NNTPLVKAS-----GRELCYILLFGVLSLYCYMTFFIAKPSVICALRRRLGLGSTFAIC 655

QY 131 FSCLAHVAFNALPFR-----KNHGPRGWVI-----FTVALLLTLVEVINTENLIITLV 180
Db 656 YSALLTKT---NCIARIFDGVRNGAQRPFISPSQVFICLGLILVQIVMVSWMLILETP 712

QY 181 RGSSEGGPOGNSSAGNAVAPSPCAIANMDFVMALIVYMLLLGAFLGAWPALCGRYK-RWR 239
Db 713 GTRYTYLPEKRET----VILKCNKDSSLISUTYDVOLVI-----LCTVYAFKR 759

QY 240 K-----HGVFLLTTATTSVAIWVWMVITYGKNKHNSPTWDPTLATALAANAWAVYL 293
Db 760 KCPENFEAKFIQFTWTTCIIWLALPIFYVNTSSDYRVQT---TTMCISVSLSGFVVLG 816

QY 294 FYVIPEV 300
Db 817 CLFAPKV 823

RESULT 4
A36455
A:Title: Induction of cell fate in the Drosophila retina: the bride of sevenless prot
A:Reference number: A36455; MUID:91115074
A:Accession: A36455
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-896 <HA2>
A:Cross-references: GB:X55887
R:Hart, A.C.; Harrison, S.D.; Van Vactor Jr., D.L.; Rubin, G.M.; Zipursky, S.L.
Proc. Natl. Acad. Sci. U.S.A. 90, 5047-5051, 1993
A:Title: The interaction of bride of sevenless with sevenless is conserved between Dr
A:Reference number: A47550; MUID:93281693
A:Accession: A47550
A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
A:Molecule type: DNA
A:Residues: 4-896 <HA2>
A:Cross-references: GB:L08133
C:Genetics:
A:Gene: FlyBase:boss
A:Cross-references: FlyBase:FBgn0000206
A:Introns: 18/1: 175/1: 325/1: 811/3
C;Keywords: transmembrane protein

Query Match 6.8%; Score 159; DB 2; Length 896;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2002, 09:29:21 : Search time 20.9 seconds
(without alignments)
2027.531 Million cell updates/sec

Title: US-09-895-686-1
Perfect score: 2326
Sequence: 1 MAHKALVMCLGLPLFLFP.....ATPKDGKNSQVFRNPYWD 441

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues
Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	167	7.2	1099	2 T16283	hypothetical prote
2	160.5	6.9	879	2 JH0562	metabotropic gluta
3	160.5	6.9	879	2 JC7160	metabotropic gluta
4	159	6.8	896	2 A36455	bride of sevenless
5	157	6.7	896	2 S26740	gene boss protein
6	152	6.5	872	2 JH0561	metabotropic gluta
7	147	6.3	999	2 T27628	hypothetical prote
8	146	6.3	893	2 A47550	bride of sevenless
9	142	6.1	1199	2 A41939	G protein-coupled
10	133	5.7	1218	2 S71376	glutamate receptor
11	132.5	5.7	551	2 T30806	metabotropic gluta
12	125	5.4	908	2 I49142	metabotropic gluta
13	118.5	5.1	1079	2 I59362	calcium/polyvalent
14	118	5.1	1085	2 S40476	Ca(2+)-sensing rec
15	117.5	5.1	699	2 H87275	thio-disulfide int
16	117.5	5.1	1078	2 A56715	calcium receptor (
17	117.5	5.1	1088	2 B56715	calcium receptor (
18	112.5	4.8	912	2 JH0563	metabotropic gluta
19	110.5	4.8	871	2 A48742	metabotropic gluta
20	108	4.6	347	2 G89003	protein T24A6.6 [i
21	108	4.6	487	2 S22350	secy protein - Hal
22	107	4.6	477	2 AB2866	excoq-like protein
23	107	4.6	480	2 H97642	excoq-like protein
24	107	4.6	646	2 AC1339	ABC transporter (p
25	106.5	4.6	477	2 D70546	hypothetical prote
26	106.5	4.6	494	2 B89827	hypothetical prote
27	106.5	4.6	642	2 H83743	ABC transporter (p
28	106	4.6	409	2 E97993	cell division prot
29	106	4.6	463	2 AC0281	probable amino aci

ALIGNMENTS

RESULT 1

T16283

hypothetical protein F35H10.10 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C:Accession: T16283

R:Woessner, J.

submitted to the EMBL Data Library, November 1995

A:Description: The sequence of C. elegans cosmid F35H10.

A:Reference number: Z18490

A:Accession: T16283

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-1099 <WOE>

A:Cross-references: EMBL:U40934; NID:g1072149; PID:g1072155; PIDN:AAA81683.1; CESP:F3

C:Genetics:

A:Gene: CESP-F35H10.10

A:Introns: 35/3; 214/3; 548/1; 666/3; 740/1; 798/1; 869/2; 984/1; 1059/3

Query Match

Best Local Similarity 7.2%; Score 167; DB 2; Length 1099;

Matches 70; Conservative 55; Mismatches 140; Indels 72; Gaps 11;

QY	46	SGAWG-----IVLEAVAGAGIVTFTVLITILVASLPFVQDTKRSLLGTQVFLGLTLC	99
DB	710	TGHRDHPHNYVLLALITLVVVAIAVLVLVKL--YLRVVRGNSLG--ISLLIGIII	765
QY	100	LFCLVFACVVKPDFSTCASRFLGVLFAICFSCLAHVAFALNFKNNHGRGWIFTV	159
DB	766	LYSTAFFFPDPTDSCRLRVILHGLGYTICFGVMIKAKATQLRNAETLGGF-----	816
QY	160	ALLTLVEIINTEWLIITLIVRGSGEGPQGNSSAGWAV-----ASPCA	204
DB	817	---TAHISFNWYLLLEFI-----VGQIALSISWLEFPFMSTICVIDTNVQRMMCTM	867
QY	205	ANMDFVALIYVMLLGLGAPLGPANPALCGR-YRRKRKHGVFVLLTTATSVAINVWVIMY	263
DB	868	GKVEFVSVFYVMILI---FMALFISMLNRNKRNYKETKWLLYSTVLCFTTVAWITLY	924
QY	264	TYGNKQHNSTWDDPTLATAAANAFVLFYVIVEPSQVTKSSP-EOSYOGDMYPTRGV	322
DB	925	LVLDHE-----FRDTVIVVELVACATILGLFGPKIYIILLSEVPVVAFKRPFEPN---	976
QY	323	GYETILKEQKQSMFVENKAFSDPEPVAARPPVSPYS	359
DB	977	-----HTDLFEKDDDLPSQRAVSPAS	997

RESULT 2

JH0562

metabotropic glutamate receptor 3 precursor - rat

```
QY 102 CLVFACVVKPDEFSTCASRRRLFGVLFACFSCLAHV--FALNFLAR-KNHGPRGW---- 154
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
663 SSSLEFFIGEPQDWTCRLRQPAFGISFVLCISILVKTNRVLLVFEAKIPTSFHRKMWGLN 722
QY 155 VIFTVALLLLTLVEVIINTEWLIITLVRGSGEGGPGQGNSSAGWAVASPCAIANMDFVMALI 214
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
723 LQFLLVFLCTFMQILICIIWLY-----TAPSSYRNHELEDEII 761
QY 215 YV-----MLLLGAFLG---AWPALCGRYK-RWRK-----HGVEVLLTTTATSVAIWVVW 259
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
762 FITCHEGSLMALGSLIGYTCLLAAICEFFPAFKSKKLPENFENAKFITFSMLIFFIWIWISF 821
QY 260 IVMY--TYG 266
Db : : : : :
822 IPAYASTYG 830
```

Search completed: September 16, 2002, 09:36:53
Job time: 222 sec


```
FT CARBOHYD 480 480 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 565 565 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 908 AA; 101866 MW; FCAB54CB8E3DD915 CRC64;

Query Match 5.2%; Score 122; DB 1; Length 908;
Best Local Similarity 20.8%; Pred. No. 0.22;
Matches 75; Conservative 70; Mismatches 129; Indels 86; Gaps 22;

QY 49 WGIIVLEAVAGAGIV-TTEVL-TIILVASLPFVODTKRSLSLGTVQVFLGLTGLF---CL 103
   I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db 583 WAVVPFIALGIIATTFVIVTRNDPIVRAS-----GRELSVLLT-GIFLCYSI 635
   I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :

QY 104 VFACVVKPFSTCASRFLFGLFAICFCLAARHVFALNFAR-----KNHGP 151
   I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db 636 TFLMIAAPDIICSFRRIFLGL-GMCF5-YAALLTKNRIHFEGGKKSVTAPRISP 692
   I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :

QY 152 RGVWIFTVALL-ITLVEVIINTWLLI-----TLVRSGEGGPGNCSAGWAVASPAIAN 206
   : : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db 693 ASQLVITFSLISVQLLGVFV---WFVVDPPTHTIDYGEQRTLPENARG---VLKCDISD 746
   : : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :

QY 207 MDFWMLIVVWLLLGAFLGAWPALCGRYKRWKHGV-----FVLLTATSVAI 255
   : : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db 747 LSLICSLGYSILLMV-----TCTVY-AIKTRGVPTFNEAKPIGFTMYTTC---II 793
   : : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :

QY 256 WYVWIVMYTYGNKQHSPTW-DDPTLALALANAFVLFYVPEVSQVTKSSPEQSYQG 314
   I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db 794 WLAIFIPF-PGTAQSAEMKYIOTTLLVSMSLSASVGLMYPKV-YIIFIPEQNVQ- 850
   I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :

QY 315 DMYPTRGVGYETILKEQKQSGMVENKAFSMDPEVAAKRPVSYGVNGLLTYSVYOPT 374
   I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db 851 ----KKRSEKAVVTAATMQSKLIQK---GNDRP-----NGEVKSELCELSLE 890
   I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :

RESULT 14
CASR_MOUSE
ID CASR_MOUSE STANDARD; PRT; 1079 AA.
AC Q9QY96; O089519; Q9QY95; Q9QZ08; Q9R1D6; Q9R1Y2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Extracellular calcium-sensing receptor precursor (CaSR) (Parathyroid
GN Cell calcium-sensing receptor).
DE CASR OR GPRC2A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.
RP STRAIN=C57BL/6; TISSUE=Kidney;
RX MEDLINE=20092890; PubMed=10625662;
RA Oda Y., Tu C.-L., Chang W., Crumrine D., Koemueves L., Mauro T.,
RA Elias P.M., Bikle D.D.;
RT "The calcium sensing receptor and its alternatively spliced form in
RT murine epidermal differentiation.";
RL J. Biol. Chem. 275:1183-1190(2000).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM A), AND TISSUE SPECIFICITY.
RX STRAIN=BLACK SWISS X 129/SVJ; TISSUE=Kidney;
RX MEDLINE=20119279; PubMed=10652312;
RA Pi M., Garner S.C., Flannery P., Spurney R.F., Quarles L.D.;
RT "Sensing of extracellular cations in Casr-deficient osteoblasts.
RT Evidence for a novel cation-sensing mechanism.";
RL J. Biol. Chem. 275:3256-3263(2000).
RN [3]
RP SEQUENCE OF 256-600 FROM N.A. (ISOFORM A), AND TISSUE SPECIFICITY.
RC TISSUE=Epiphyseal cartilage;
RX MEDLINE=20043955; PubMed=10579354;
RA Chang W., Tu C., Chen T.-H., Komuves L., Oda Y., Pratt S.A.,
RA Miller S., Shoback D.;
RT "Expression and signal transduction of calcium-sensing receptors in
RT cartilage and bone.";
RN [4]
RP Endocrinology 140:5883-5893(1999).
RA SEQUENCE OF 507-582 FROM N.A. (ISOFORM A).
RC STRAIN=NMRI; TISSUE=Brain;
RA Hildenbrand J., Ammon H.P.T., Wahl M.A.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 562-814 FROM N.A.
RC TISSUE=Kidney;
RA Moawad T.I., Riccardi D.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE OF 646-799 FROM N.A., AND TISSUE SPECIFICITY.
RX MEDLINE=97231187; PubMed=9076582;
RA Quarles L.D., Hartle J.E. II, Siddhanti S.R., Guo R., Hinson T.K.;
RT "A distinct cation-sensing mechanism in MC3T3-E1 osteoblasts
RT functionally related to the calcium receptor.";
RL J. Bone Miner. Res. 12:393-402(1997).
CC -!- FUNCTION: SENSES CHANGES IN THE EXTRACELLULAR CONCENTRATION OF
CC CALCIUM IONS. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY A
CC G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
CC MESSENGER SYSTEM.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A (SHOWN HERE) AND B; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: EPIDERMIS, KIDNEY AND CARTILAGE.
CC -!- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF110178; AAD28371.1; -
DR EMBL; AF110179; AAD28372.1; -
DR EMBL; AF128842; AAD40638.1; -
DR EMBL; AF068900; AAC19388.1; -
DR EMBL; AB027140; BAA77688.1; -
DR EMBL; AF002015; AAC53252.1; -
DR EMBL; AF159565; AAF00193.1; -
DR MGD; MGI:1351351; Casr.
DR InterPro; IPR001828; ANF_receptor.
DR InterPro; IPR000337; GPCR_Mgr.
DR Pfam; PF00003; 7tm_3; 1.
DR Pfam; PF01094; ANF_receptor; 1.
DR PRINTS; PR00248; GPCRMR.
DR PROSITE; PS00979; G_PROTEIN_RECEP_F3_1; 1.
DR PROSITE; PS00980; G_PROTEIN_RECEP_F3_2; 1.
DR PROSITE; PS00981; G_PROTEIN_RECEP_F3_3; 1.
DR PROSITE; PS50259; G_PROTEIN_RECEP_F3_4; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
FT SIGNAL 1 19
FT CHAIN 20 1079
FT DOMAIN 20 612
FT TRANSMEM 613 635
FT TRANSMEM 636 649
FT TRANSMEM 650 670
FT TRANSMEM 671 681
FT TRANSMEM 682 700
FT DOMAIN 701 724
FT TRANSMEM 725 745
FT DOMAIN 746 769
FT TRANSMEM 770 792
FT DOMAIN 793 805
FT TRANSMEM 806 828
FT TRANSMEM 829 836
FT TRANSMEM 837 862
FT TRANSMEM 863 1079
FT CARBOHYD 90
   N-LINKED (GLCNAC. . .) (POTENTIAL).
```


Tanabe Y., Masu M., Ishii T., Shigemoto R., Nakanishi S.;
 "A family of metabotropic glutamate receptors.";
 Neuron 8:169-179(1992).
 [4]
 ALTERNATIVE SPLICING (ISOFORM 1C).
 TISSUE=Brain;
 MEDLINE=93066232; PubMed=1438218;
 Pin J.-P., Waeber C., Prezeau L., Bockaert J., Heinemann S.F.;
 "Alternative splicing generates metabotropic glutamate receptors
 inducing different patterns of calcium release in Xenopus oocytes.";
 Proc. Natl. Acad. Sci. U.S.A. 89:10331-10335(1992).
 CC -!- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR IS
 MEDIATED BY A G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-
 CALCIUM SECOND MESSENGER SYSTEM. MAY PARTICIPATE IN THE CENTRAL
 ACTION OF GLUTAMATE IN THE CNS. SUCH AS LONG-TERM POTENTIATION IN
 THE HIPPOCAMPUS AND LONG-TERM DEPRESSION IN THE CEREBELLUM.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1A (SHOWN HERE), 1B AND 1C;
 ARE PRODUCED BY ALTERNATIVE SPLICING. 1B AND 1C ARE C-TERMINALLY
 TRUNCATED FORMS OF 1A.
 CC -!- TISSUE SPECIFICITY: IS PREDOMINANTLY EXPRESSED IN CEREBELLAR
 PURKINE CELLS, CA2-CA3 PYRAMIDAL CELLS OF THE HIPPOCAMPUS, AND
 MITRAL AND TUFTED CELLS OF THE OLFACTORY BULB.
 CC -!- MISCELLANEOUS: ACTIVATED BY QUISQUALATE > GLUTAMATE > IBOTENATE >
 TRANS-1- AMINOCYCLOPENTYL-1,3-DICARBOXYLATE; INHIBITED BY
 2-AMINO- 3-PHOSPHONOPROPIONATE.
 CC -!- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC STRONGEST, TO MGLUR8.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X57569; AAA40799.1; -;
 DR EMBL; M61099; AAB19497.1; -;
 DR EMBL; S48085; AAB24138.1; -;
 DR PIR; S15362; S15362.
 DR PIR; A41939; A41939.
 DR GCRdb; GCR_0216; -;
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam; PF00003; 7tm_3; 1.
 DR Pfam; PF01094; ANF_receptor; 1.
 DR PRINTS; PR00248; GPCRMR.
 DR PROSITE; PS00979; G_PROTEIN_RECP_F3_1; 1.
 DR PROSITE; PS00980; G_PROTEIN_RECP_F3_2; 1.
 DR PROSITE; PS00981; G_PROTEIN_RECP_F3_3; 1.
 DR PROSITE; PS00981; G_PROTEIN_RECP_F3_4; 1.
 DR PROSITE; PS02599; G_PROTEIN_RECP_F3_4; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Multigene family; Alternative splicing.
 FT SIGNAL 1 18 POTENTIAL.
 FT CHAIN 19 1199 METABOTROPIC GLUTAMATE RECEPTOR 1.
 FT DOMAIN 19 592 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 593 615 I (POTENTIAL).
 FT DOMAIN 616 629 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 630 650 II (POTENTIAL).
 FT DOMAIN 651 661 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 662 680 III (POTENTIAL).
 FT DOMAIN 681 706 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 707 727 IV (POTENTIAL).
 FT DOMAIN 728 750 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 751 772 V (POTENTIAL).
 FT DOMAIN 773 785 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 786 808 VI (POTENTIAL).
 FT DOMAIN 809 814 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 815 840 VII (POTENTIAL).
 FT DOMAIN 841 1199 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 1014 1034 GLN/PRO-RICH.
 FT DOMAIN 1074 1080 GLN/PRO-RICH.

FT	DOMAIN	1126	1135	ASP/GLU-RICH (ACIDIC).
FT	DOMAIN	1140	1199	SER-RICH.
FT	CARBOHYD	98	98	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	223	223	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	397	397	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	515	515	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARSPLIC	887	906	NNSGKSVSWSEPGGROAPKG -> KKRQPEFSPSQPSAH
FT	VARSPLIC	907	1199	AOL (IN ISOFORM 1B).
FT	VARSPLIC	888	897	MISSING (IN ISOFORM 1B).
FT	VARSPLIC	898	1199	SNKGSVSWSE -> FALDRQNTVY (IN ISOFORM 1C).
FT	VARSPLIC	898	1199	MISSING (IN ISOFORM 1C).
SQ	SEQUENCE	1199 AA;	133235 MW;	EEE5A04C50694B9F CRC64;

Query Match 6.1%; Score 142; DB 1; Length 1199;
 Best Local Similarity 22.4%; Pred. No. 0.011;
 Matches 85; Conservative 54; Mismatches 153; Indels 88; Gaps 19;

QY	20	GAWAQGHVPCCSQGLNPLVYNLCDSRGANGIVLEAVAGAGIVTFTVLTILV--ASLPF 77
DB	568	GWPNAEL-----TGCEPIPVRYLEWSDIEIIAIFSCGLILVTLFVTLFVLYRDTVP 622
QY	78	VODTKRSLLGTVFFLLGT-LGLFCLVFACVVKVPDFSTCASRRFLFGVLFACFSCIAA 136
DB	623	VKSSRELC-----YIILAGIFLGVC-PFTLIAKPTTSCYLQRLVGLSSAMCYSAL-- 675
QY	137	HVFALNPLARKNUG-----PR---GWVIFTVALLLTLVEVINTWELIITLVKRG 184
DB	676	-VTKTRIRIARISAGKKIKCTRKPRFMSAWAQVILISIVOLT-----LVVTLI---- 725
QY	185	EGGPGQGNSSAGWAVASP-----CATANMDFVMAIIVMVLILLGLAFLGAW 228
DB	726	-----IMPEPMPILSPYSIKEVYLICNTSNLGVAPVGVNGLLISCIYYAF 772
QY	229	-----PALCGRYKRWKRGVFLVLLTTATSVAIWVWVWVYTYGNKHQNSPTDPTLAI 283
DB	773	KTRNVPANFEAK-----YIAFTMYTTCIWLAFVPIY-FGSNYSKIITCFVAVLSVT 824
QY	284	LAANAFVLFYVI---PE---VSQVTKSSPEQSYOGD-MYPTRGVGYETILKEQK-QQS 335
DB	825	VALGCMFTPKMYIIIAKPERNVSFAETTSQVDMHMGVDGKLPKCRSTNLTFLNIFRKRKPG 884
QY	336	MEVEN-KAPSMDEPVAAKRP 354
DB	885	NANSNGKSVSWSEPGGROAP 904

RESULT 11
 MGR8_MOUSE
 ID MGR8_MOUSE STANDARD; PRT; 908 AA.
 AC P47743;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Metabotropic glutamate receptor 8 precursor.
 GN GRM8 OR GPRC1H OR MGLUR8.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Retina;
 RX MEDLINE=95239344; PubMed=7722646;
 RA Duvoisin R.M., Zhang C., Ramonell K.;
 RT "A novel metabotropic glutamate receptor expressed in the retina and
 olfactory bulb.";
 RL J. Neurosci. 15:3075-3083(1995).
 CC -!- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR
 IS MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLATE CYCLASE
 CC ACTIVITY.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: STRONGLY EXPRESSED IN OLFACTORY BULB,


```
Db 549 SLTGFCFPEQYIRWGDAMAVGPVTIAC-----LGA-LAT 582
QY 64 TFVLITILVASLPFVQDTRKRSILGTQVFFELTGL-FCLVFACVVKPDEFSCASRRFL 122
Db 583 LFLVGVFVRHNATPVVKASGREL-----CYILGGVFLCYCMTEIFTAKSTAVCTIRRLG 638
QY 123 FGVLFACFSCLAHFAVNLFLARKNHGPRGW-----IFTVALLLTLV--EVIINT 172
Db 639 LGTAFSVCYSAALLTKT---NRIARIFGAREGAORPREFISPAQVAICLALISGQLLIV 695
QY 173 EWLIIILVRSGEGGQGNSSAGWAVASCAIANMDFVVALIYVMLLLGAFGLANPALC 232
Db 696 AMLVLRAPGTGKETAPERRE-----VTLRCNHRDASHMLGSLAYNLLI-----ALC 742
QY 233 GRYK-RWRK-----HGVFVLLTTATSVAIWVVMVMTYGNKQHNPTWDDPTLAIALA 285
Db 743 TLAFNTRKCPENFNEAKFIFGFTMYTTCILWALLPFIYVTSDDYRQT---TTMCVSVS 799
QY 286 ANAWAFVLYVPEVSQVTKSSPEQSQGDMDYPTRGV 323
Db 800 LSG-SVVLGCLFAPKHLIILFQPKNVVSHRAPTSRFG 836

RESULT 9
BOSS_DROVI STANDARD; PRT; 893 AA.
AC Q24738;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Bride of sevenless protein precursor.
GN BOSS.
OS Drosophila virilis (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7244;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93281693; PubMed=8506350;
RA Hart A.C., Harrison S.D., van Vactor D.L. Jr., Rubin G.M.,
RA Zipursky S.L.;
RT "The interaction of bride of sevenless with sevenless is conserved
RL between Drosophila virilis and Drosophila melanogaster.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:5047-5051(1993).
CC -!- FUNCTION: ACTS AS A LIGAND FOR SEVENLESS TYROSINE-KINASE RECEPTOR
CC DURING EYE DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: L08132; AAA72332.1;
DR FlyBase; FBgn0013105; Dvir\boss.
DR InterPro; IPR002956; Bride_of_7less.
DR InterPro; IPR000337; GPCR_Mgr.
DR Pfam; PF00003; 7tm_3; 1.
DR PRINTS; PR01223; BRIDOF7LESS.
DR Transmembrane; Glycoprotein; Vision; Signal.
KW SIGNAL
FT CHAIN 1 30 POTENTIAL.
FT CHAIN 31 893 BRIDE OF SEVENLESS PROTEIN.
FT DOMAIN 32 530 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 531 551 POTENTIAL.
FT TRANSMEM 563 583 POTENTIAL.
FT TRANSMEM 607 627 POTENTIAL.
FT TRANSMEM 630 650 POTENTIAL.
FT TRANSMEM 653 673 POTENTIAL.
FT TRANSMEM 692 712 POTENTIAL.
```

```
FT TRANSMEM 722 742 POTENTIAL.
FT TRANSMEM 772 893 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 773 893 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 183 183 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 307 307 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 328 328 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 471 471 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 482 482 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 893 AA; 99939 MW; ABB9DAC21DB680C3 CRC64;

Query Match 6.3%; Score 146; DB 1; Length 893;
Best Local Similarity 20.8%; Pred. No. 0.0043;
Matches 71; Conservative 58; Mismatches 136; Indels 76; Gaps 15;

QY 34 GLNPLYNCLDRSGA--WGIVLEAVAGAGIVTTFVLITILVASLPFVQDTRKRSLLG--- 88
Db 503 GENRRYPFLFDGESVWFRIKLDTWATGLTAAILGLIATLAILVFI---VVRISLGDVF 559
QY 89 -----TQVFFELTGLTGLFC--LVFAC-----VVKPDFST---CASRRFLFCVLV 127
Db 560 EGNPVTISILLLSILVFCSEVFPFSEYVGEQRNSHVTFEDVHTLNTLCGVRVFIITLVY 619
QY 128 AICFSCIAAHVFAL-----NFLARKNHGPRGWITFTVALLTLVEVIINTWLLITLV 180
Db 620 CFVFSLLLCRAVMLASIGSEGGFLSHVN---GYIOAIIICVLSVFVQVGMVQLLVV--M 673
QY 181 RGSSEGGPOGNSAG--WAVASPCAIAANMDFVVALIYVMLLLGAFGLANPALCGRYKWR 239
Db 674 HLASESVSCENIYYGRW-----LWGLLAYDFLLLCSLVSLVFPFYRSQRNRY 720
QY 240 KHGVFVLLTTATSVAIWVVMVMTYGNKQHNPTWDDPTLAIALAANAWAFVLYVIP- 298
Db 721 E-GILIVIGAVLLIINSWIAMSLEFDE-----WRDAIPLGMQASGAWLVGLILPR 773
QY 299 -----EVSQVTKSSPEQS-----YQGDMDYPTRGVGYETI 327
Db 774 TFLIVRGIERSDIAQALPSLTSLAFAQNNOYSSEQSVYECV 814

RESULT 10
MGRL_RAT STANDARD; PRT; 1199 AA.
AC P23385;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Metabotropic glutamate receptor 1 precursor.
GN GRM1 OR GPRC1A OR MGLUR1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Brain;
RX MEDLINE=91156047; PubMed=1847995;
RA Masu M., Tanabe Y., Tsuchida K., Shigemoto R., Nakanishi S.;
RT "Sequence and expression of a metabotropic glutamate receptor.";
RL Nature 349:760-765(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=92022526; PubMed=1656524;
RA Houamed K.M., Kuijper J.L., Gilbert T.L., Haldeman B.A., O'Hara P.J.,
RA Mulvihill E.R., Almers W., Hagen F.S.;
RT "Cloning, expression, and gene structure of a G protein-coupled
RL glutamate receptor from rat brain.";
RL Science 252:1318-1321(1991).
RN [3]
RP ALTERNATIVE SPLICING (ISOFORM 1B).
RC TISSUE=Brain;
RX MEDLINE=92110002; PubMed=1309649;
```


CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch).

```

DR EMBL: Z47073: CAA07374.1; -.
DR WormPep: ZC506.4; CE01682.
DR InterPro: IPR001828; ANF_receptor.
DR InterPro: IPR000337; GPCR_Mgr.
DR Pfam: PF00003; 7tm_3; 1.
DR Pfam: PF01094; ANF_receptor; 1.
DR PRINTS: PR00248; GPCRMR.
DR PROSITE: PS00979; G_PROTEIN_RECEP_F3_1; 1.
DR PROSITE: PS00980; G_PROTEIN_RECEP_F3_2; 1.
DR PROSITE: PS00981; G_PROTEIN_RECEP_F3_3; 1.
DR PROSITE: PS0259; G_PROTEIN_RECEP_F3_4; 1.
DR KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT TRANSMEM 682 704 POTENTIAL.
FT TRANSMEM 719 739 POTENTIAL.
FT TRANSMEM 751 769 POTENTIAL.
FT TRANSMEM 792 812 POTENTIAL.
FT TRANSMEM 836 857 POTENTIAL.
FT TRANSMEM 871 893 POTENTIAL.
FT TRANSMEM 904 929 POTENTIAL.
FT DOMAIN 888 891 POLY-PHE.
FT CARBOHYD 518 518 N-LINKED (GLCNAC..) (POTENTIAL).
SQ SEQUENCE 999 AA; 113275 MW; AB22AF2A28D9AOCF CRC64;

```

Qy	30	GCSOGLNPLIYNLCDSRGANGIVLE-----AVAGAGIVTTFVLITIL	71		
Dd	650	CGEQG-----WMPTKDRKGCFDLSQLKYMWRWSMYSLVPTILAFVGI-IATLFVIVVVV	704		
Qy	72	VASLPFVDQTKKRLLGTQVFFFLTGLGCLVEACVVKPDFSTCASRRFLFGVLFAICF	131		
Dd	705	IYNETPVVKASGREL---SYILLISMIMCYMTFVLLSKPSAIVCAIKR--TGIGFA--F	757		
Qy	132	SCLAHHVPALN-----FLARKNHGR-----GWIFTVALLLLDELVEVIINTEMLIITLV	180		
Dd	758	SCLYSAMFEVTNRIFRIEFTSRSAQRPRFISPISQVMATA--MLAGVOLIGLSLWLSVV--	813		
Qy	181	RSGEGGPQGNSSAGW-----AVASPCAIAINMDFMALIIYVMLLLLGAFLGAWPALC	232		
Dd	814	-----PGCRHHPHYTRQOVVLTCNVPDHFELYSLAYDGFLV-----LC	852		
Qy	233	GRYK-RWRK-----HGVEVLLTTATSVAIWWMVIMVITYKNKHNSPWTDDPTLAIALA	285		
Dd	853	TTVAKTRKPENETNKFKIGFSMYTTCVWLMSWIFFEF-FCGSDFOQTOTSSLGCISISMS	911		
Qy	286	ANAAFWLFVYIPEVSQVTRKSPE-----OSYGDMYPTRGVGYETILKBOKGQSMEVE	339		
Dd	912	ANVALACIF-----SPKLWIILFEKHN-----VRQEGESMLNK	946		
Qy	340	NK-----AFSMDPEPVAARPPSPSYGGYNGOLLTSVYTQMAMLMHKHVSEGAJD	388		
Dd	947	SSRLGNCSSRLCANSIDEP-----NOYTALLTDSTRRSRKRTSOPTSTSSAHD	996		
Qy	389	IIL 391			
Dd	997	TFL 999			
RESULT 7					
MGRL_HUMAN					
ID	MGRL_HUMAN	STANDARD;	PRT; 1194 AA.		
AC	Q13256; Q14758;				
DT	01-NOV-1997 (Rel. 35, Created)				
DT	01-NOV-1997 (Rel. 35, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Metabotropic glutamate receptor 1 precursor.				
GN	GM1 OR GPRC1A OR MGLUR1.				

Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
MEDLINE=96029774; PubMed=7476890;
Desai M.A., Burnett J.P., Mayne N.G., Schoepp D.D.;
RT "Cloning and expression of a human metabotropic glutamate receptor 1
alpha; enhanced coupling on co-transfection with a glutamate
transporter.";
RT Mol. Pharmacol. 48:648-657(1995).
RL [2]
RN
SEQUENCE FROM N.A.
MEDLINE=97231349; PubMed=9076744;
Stephan D., Bon C., Holzwarth J.A., Galvan M., Pruss R.M.;
RA "Human metabotropic glutamate receptor 1: mRNA distribution,
RT chromosome localization and functional expression of two splice
RT variants.";
RL Neuropharmacology 35:1649-1660(1996).
RN
-!- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR IS
CC MEDIATED BY A G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-
CC CALCIUM SECOND MESSENGER SYSTEM. MAY PARTICIPATE IN THE CENTRAL
CC ACTION OF GLUTAMATE IN THE CNS, SUCH AS LONG-TERM POTENTIATION IN
CC THE HIPPOCAMPUS AND LONG-TERM DEPRESSION IN THE CEREBELLUM.
CC
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; ALPHA (SHOWN HERE) AND BETA;
CC ARE PRODUCED BY ALTERNATIVE SPLICING.
CC
CC -!- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
CC STRONGEST, TO MGLUR5.
CC
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement>
CC or send an email to license@isb-sib.ch).
CC
CC
CC EMBL: U31215; AAA87843.1; -;
DR EMBL: U31216; AAA87844.1; -;
DR EMBL: L76627; AAB05337.1; -;
DR EMBL: L76631; AAB05338.1; -;
DR GCRDB: GCR_1825; -;
DR GCRDB: GCR_1826; -;
DR GCRDB: GCR_1982; -;
DR GCRDB: GCR_1983; -;
DR MIM: 604473; -;
DR InterPro: IPR001828; ANF_receptor.
DR InterPro: IPR000337; GPCR_Mgr.
DR Pfam: PF00003; 7tm_3; 1.
DR Pfam: PF01094; ANF_receptor; 1.
DR PRINTS: PR00248; GPCRMRG.
DR PROSITE: PS00979; G_PROTEIN_RECPEP_F3_1; 1.
DR PROSITE: PS00980; G_PROTEIN_RECPEP_F3_2; 1.
DR PROSITE: PS00981; G_PROTEIN_RECPEP_F3_3; 1.
DR PROSITE: PS02059; G_PROTEIN_RECPEP_F3_4; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Multigene family; Alternative splicing.
KW SIGNAL
FT CHAIN 1 18 POTENTIAL.
FT FT CHAIN 19 1194 METABOTROPIC GLUTAMATE RECEPTOR 1.
FT FT DOMAIN 19 592 EXTRACELLULAR (POTENTIAL).
FT FT TRANSMEM 593 615 I (POTENTIAL).
FT FT DOMAIN 616 629 CYTOPLASMIC (POTENTIAL).
FT FT TRANSMEM 630 650 II (POTENTIAL).
FT FT DOMAIN 651 661 EXTRACELLULAR (POTENTIAL).
FT FT TRANSMEM 662 680 III (POTENTIAL).
FT FT DOMAIN 681 706 CYTOPLASMIC (POTENTIAL).
FT FT TRANSMEM 707 727 IV (POTENTIAL).
FT FT DOMAIN 728 750 EXTRACELLULAR (POTENTIAL).
FT FT TRANSMEM 751 772 V (POTENTIAL).
FT FT DOMAIN 773 785 CYTOPLASMIC (POTENTIAL).
FT FT TRANSMEM 786 808 VI (POTENTIAL).

```

QY 123 FGVLEAIFCSLAHVAFLNFKNHGPRGW-----IFTVALLLFLV---EVIINT 172
Db LGTAFSVCYSSALLTKT---NRIARIFGAREGAQRPFISPAQVAICLALISSQLLIVA 695
QY 173 EWLIITLVROSGGGGQGNSSAGWASVAPCAIANMDFVMAIIVVMLLLGAFLGAWPALC 232
Db 696 AYLWVEAPGKGKETAPERRE---VVTLRCHNRDASMLGSLAYNLLI-----ALC 742
QY 233 GRYK-RWRK-----HGVFVLLTATSAVAIVVWVYTYGNKOHNSPTWDDPTLALALA 285
Db 743 TLXAFTRKCPENFNEAKFTGFTMYTTCIIWLAFPLIFVYVTSDDYRVQT---TTCMVSVS 799
QY 286 ANAWAFVLFVIVPEVSVOTKSSPEQSYOGDMYPTRGV 323
Db 800 LSG-SVVLGCLFAPKLHILFQPKNVVSHRAPTSREG 836

RESULT 5
MGR3_HUMAN
ID MGR3_HUMAN STANDARD; PRT; 877 AA.
AC Q14832;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Metabotropic glutamate receptor 3 precursor.
GN GRM3 OR GPRC1C OR MGLUR3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=96437205; PubMed=8840013;
RA Makoff A., Voipe F., Ielchuk R., Harrington K., Emson P.;
RT "Molecular characterization and localization of human metabotropic
RT glutamate receptor type 3.";
RL Brain Res. Mol. Brain Res. 40:55-63(1996).
CC -!- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR
CC IS MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLATE CYCLASE
CC ACTIVITY.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
CC STRONGEST, TO MGLUR2.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X77748; CAA54796.1; -.
DR GCRDB; GCR_2070; -.
DR MTM; 601115; -.
DR InterPro; IPR001828; ANF_receptor.
DR InterPro; IPR000337; GPCR_Mgr.
DR Pfam; PF000003; 7tm_3; 1.
DR Pfam; PF01094; ANF_receptor; 1.
DR PRINTS; PR00248; GPCRMRG.
DR PROSITE; PS00979; G_PROTEIN_RECEP_F3_1; 1.
DR PROSITE; PS00980; G_PROTEIN_RECEP_F3_2; 1.
DR PROSITE; PS00981; G_PROTEIN_RECEP_F3_3; 1.
DR PROSITE; PS0259; G_PROTEIN_RECEP_F3_4; 1.
DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Multigene family.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 877 METABOTROPIC GLUTAMATE RECEPTOR 3.
FT DOMAIN 21 574 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 575 597 I (POTENTIAL).
FT DOMAIN 598 611 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 612 632 II (POTENTIAL).

```

```

FT DOMAIN 633 643 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 644 662 III (POTENTIAL).
FT DOMAIN 663 686 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 687 707 IV (POTENTIAL).
FT DOMAIN 708 732 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 733 754 V (POTENTIAL).
FT DOMAIN 755 767 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 768 790 VI (POTENTIAL).
FT DOMAIN 791 800 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 801 826 VII (POTENTIAL).
FT DOMAIN 827 877 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 207 207 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 290 290 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 412 412 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 437 437 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 877 AA; 98619 MW; 66F28663CE35F740 CRC64;

Query Match 6.3%; Score 147; DB 1; Length 877;
Best Local Similarity 21.8%; Pred. No. 0.0036;
Matches 61; Conservative 55; Mismatches 106; Indels 58; Gaps 12;

QY 48 AWGIVLEAVAGIVTFTVLTIIILVA--SLPFDVQDTKRSLLTQVFFLLGTGL-LFCLV 104
Db 573 AWAIQPVTIACLGFMCTCMVTVFIKHNTPLVKASGRE-----LCYILLFGVGLSYCMT 627
QY 105 FACVVRPDFSTCASRRFLGVLEAIFCSLAHVAFLNFKLAR-----KNHGPRGWVI--- 156
Db 628 FFFIAKPSVPICALRLGLGSSFAICYSAULTTKT---NCIARIFDGVKNGAQRKPFISPS 684
QY 157 --FTVALLTLVEIINTEMLIITLVRGSGEGPGQGNSSAGWAVAS-----PCAIANM 207
Db 685 SQVFICGLGLIIVQVWVWLLI-----EAPTRRYTLAEKRETVILKCNVKDS 733
QY 208 DFVMAIYVMLLLGLAFGLGAWPALCGRYK-RWRK-----HGVFVLLTATSAVAIVVWVI 260
Db 734 SMLISLTYDVILVI-----LCTVVAFKTRKCPENFNEAKFIGTMTVTTCTIWLAF 784
QY 261 VMITYGNKHNSPTWDDPTLALALANAWAFVLFYVIPEV 300
Db 785 PIFVYTSDDYRVQT---TTMCISVSGVGVVGLCLFAPKV 821

RESULT 6
MGR1_CAEL
ID MGR1_CAEL STANDARD; PRT; 999 AA.
AC Q09630;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Probable metabotropic glutamate receptor mgl-1.
GN MGL-1 OR ZC506.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Harris B.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR
CC IS MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLATE CYCLASE
CC ACTIVITY (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
CC STRONGEST, TO MGLUR3.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```


G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 Multigene family.
 KW SIGNAL 1 22 POTENTIAL
 FT CHAIN 23 879 METABOTROPIC GLUTAMATE RECEPTOR 3.
 FT DOMAIN 23 576 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 577 599 I (POTENTIAL).
 FT DOMAIN 600 613 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 614 634 II (POTENTIAL).
 FT DOMAIN 635 645 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 646 664 III (POTENTIAL).
 FT DOMAIN 665 688 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 689 709 IV (POTENTIAL).
 FT DOMAIN 710 734 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 735 756 V (POTENTIAL).
 FT DOMAIN 757 769 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 770 792 VI (POTENTIAL).
 FT DOMAIN 793 802 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 803 828 VII (POTENTIAL).
 FT DOMAIN 829 879 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 209 209 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 414 414 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 439 439 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 879 AA; 98959 MW; 3E5965EDD5E6DEED CRC64;
 Query Match 6.9%; Score 160.5; DB 1; Length 879;
 Best Local Similarity 22.5%; Pred. No. 0.0004;
 Matches 69; Conservative 56; Mismatches 123; Indels 59; Gaps 14;
 QY 19 PGWAQGHVPPGSGQLNPLNYNLCDR-----SGAWGIVLEAVAGAGIVTFVLTIIILVA- 73
 Db 551 PGQWPTADL-SGC-----YNLPEDYIKWEDAWAIPVTIACGLFLCTCIVITFIKH 601
 QY 74 -SLPFFVDTKKRSLGLQVFFLL--GTLGLFLVACVVKPDPSTCASRRFLFGVLFPAIC 130
 Db 602 NNTPLVKR-----GRELCYILLFGVLSYCYMTFFFIAPKSPVICALRRLGLGTSFPAIC 655
 QY 131 FSLAAHVFAFNFLAR-----KHHGPRGWI-----FTVALLTLTVEIINTENLITLV 180
 Db 656 YSALLTKT---NCIARIFDGVKNGAQRKFISSPSQVFCILGLLIVQVWVSWVLILETP 712
 QY 181 RGSGEQPGQNSAGNAVASPCAIAINDFVYALIVYMLLLIGAFLGAWPALCGRYK-RWR 239
 Db 713 GTRRYTLPEKRET---VILKCNVKDSSMLISLTVDVLI-----LCIVYAFKTR 759
 QY 240 K-----HGVFLLTATSVIAIVWVIMVYTKNQKINSPTDPTLAIALANAFVL 293
 Db 760 KCPENFNEAKFIGFTWYTTCTIWLAFILPIFYVTSSDYRVQT---TTMCISVSLSGFVVLG 816
 QY 294 FYVIPEV 300
 Db 817 CLFAPKV 823
 RESULT 3
 ID BOSS_DROME STANDARD; PRT; 896 AA.
 AC P22815; Q9VBJ5;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Bride of sevenless protein precursor.
 GN BOSS OR CG8285.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-91115074; PubMed-2276620;
 RA Hart A.C., Kraemer H., van Vector D.L. Jr., Paidhungat M.,

Zipursky S.L.;
 "Induction of cell fate in the Drosophila retina: the bride of
 sevenless protein is predicted to contain a large extracellular
 domain and seven transmembrane segments.";
 Genes Dev. 4:1835-1847(1990).
 [2]
 RP REVISIONS.
 RA Kraemer H.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE-20196006; PubMed-107311132;
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 Brannon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 Cherry J.M., Cawley S., Dahike C., Davenport L.B., Davies P.,
 de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclev J.M.,
 Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
 Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 "The genome sequence of Drosophila melanogaster";
 Science 287:2185-2195(2000).
 [4]
 RP FUNCTION.
 RX MEDLINE-91312442; PubMed-1857416;
 RA Kraemer H., Cagan R.L., Zipursky S.L.;
 RT "Interaction of bride of sevenless membrane-bound ligand and the
 sevenless tyrosine-kinase receptor.";
 RL Nature 352:207-212(1991).
 CC -!- FUNCTION: ACTS AS A LIGAND FOR SEVENLESS TYROSINE-KINASE RECEPTOR
 CC DURING EYE DEVELOPMENT.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -!- TISSUE SPECIFICITY: EXPRESSED EXCLUSIVELY BY R8 PHOTORECEPTOR
 CC CELLS AND IS INTERNALIZED IN A SEV-DEPENDENT MANNER BY R7 CELLS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X55887; CAA39373.1; -

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2002, 09:33:11 ; Search time 13.4 Seconds
(without alignments)
1274.278 Million cell updates/sec

Title: US-09-895-686-1
Perfect score: 2326
Sequence: 1 MAIHKALYVCLGLFLFPG.....ATPPKDKGNSQVFRNPYVWD 441

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	174.5	7.5	976	1 MGR_DROME	P91685 drosophila
2	160.5	6.9	879	1 MGR3_RAT	P31422 rattus norv
3	155	6.7	896	1 BOSS_DROME	P22815 drosophila
4	152	6.5	872	1 MGR2_RAT	P31421 rattus norv
5	147	6.3	877	1 MGR3_HUMAN	Q14832 homo sapien
6	147	6.3	999	1 MGR1_CAEEL	Q09630 caenorhabdi
7	147	6.3	1194	1 MGR2_HUMAN	Q13255 homo sapien
8	146	6.3	872	1 MGR2_HUMAN	Q14416 homo sapien
9	146	6.3	893	1 BOSS_DROVI	Q24738 drosophila
10	142	6.1	1199	1 MGR1_RAT	P23385 rattus norv
11	125	5.4	908	1 MGR8_MOUSE	P47743 mus musculus
12	122	5.2	908	1 MGR8_HUMAN	O00222 homo sapien
13	122	5.2	908	1 MGR8_RAT	P70579 rattus norv
14	119.5	5.1	1079	1 CASR_MOUSE	Q9qy96 mus musculus
15	118.5	5.1	1079	1 CASR_RAT	P48442 rattus norv
16	118	5.1	1085	1 CASR_BOVIN	P35384 bos taurus
17	117.5	5.1	1078	1 CASR_HUMAN	P41180 homo sapien
18	116.5	5.0	877	1 MGR6_HUMAN	O15303 homo sapien
19	112.5	4.8	912	1 MGR4_RAT	P31423 rattus norv
20	111.5	4.8	912	1 MGR4_HUMAN	Q14833 homo sapien
21	110.5	4.8	871	1 MGR6_RAT	P35349 rattus norv
22	110	4.7	1065	1 EMBB_MYCAV	P71496 mycobacteri
23	108	4.6	487	1 SECK_HALMA	P28542 haloarcula
24	106	4.6	389	1 UR2R_HUMAN	Q9ukp6 homo sapien
25	104.5	4.5	2254	1 CCAG_RAT	O54898 rattus norv
26	104	4.5	1083	1 EMBB_MYCLE	Q9cda9 mycobacteri
27	103.5	4.4	2353	1 CCAH_HUMAN	Q95180 homo sapien
28	103	4.4	416	1 RHLF_PANTR	Q28812 homo sapien
29	103	4.4	567	1 YFZ1_SCHPO	Q9us40 schizosacch
30	102	4.4	1835	1 CCAL_RAT	Q9z0y8 rattus norv
31	102	4.4	2377	1 CCAG_HUMAN	O43457 homo sapien
32	101.5	4.4	612	1 ELFI_MOUSE	O60775 mus musculus
33	101	4.3	1074	1 EMBB_MYCSM	O50393 mycobacteri

RESULT 1

MGR_DROME

ID MGR_DROME STANDARD; PRT; 976 AA.

AC P91685; Q9V485;

DT 15-JUL-1998 (Rel. 36, Created)

DT 01-MAR-2002 (Rel. 41, Last sequence update)

DT 01-MAR-2002 (Rel. 41, Last annotation update)

DE Metabotropic glutamate receptor precursor.

GN GLU-RA OR GLURA OR CG11144.

OS Drosophila melanogaster (fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RP [1]

RN SEQUENCE FROM N.A.

RC STRAIN=OREGON-R;

RX MEDLINE=96421661; PubMed=8824309;

RA Parmentier M.L., Pin J.P., Bockaert J., Grau Y.;

RT "Cloning and functional expression of a Drosophila metabotropic

glutamate receptor expressed in the embryonic CNS.;"

RL J. Neurosci. 16:6687-6694(1996).

RN [2]

RN SEQUENCE FROM N.A.

RC STRAIN=BERKELEY;

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

Ananides P.G., Scher S.E., Richards S., Ashburner M., Henderson S.N.,

George R.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

Sutton G.A., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,

Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

Abriel J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

Balwle R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,

Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

Fowler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,

Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,

Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

P49684 rattus norv
Q56072 salmonella
Q28813 pan troglod
P31533 escherichia
P26660 h genome po
Q05467 rhizobium t
Q14831 homo sapien
P35400 rattus norv
P32037 mus musculus
Q9xt50 ovic aries
P54437 bacillus su
Q60414 c ileal sod

ALIGNMENTS

Search completed: September 16, 2002, 09:36:33
Job time: 222 sec

14:

```
Db      | : | : ||          : | | | | :  
582 VFEGNPTTSLILLLSLILVFCSPVPSIEYVGGEQRNSHVTFFEDAQTLLNTLCVRVFIMTL 641
```

QY IZ8 LFAICFSCLAHVAFAL *--NFELAKKNHGFGRGWVIFIVALLILVEVINIEWLIII 179

:: ||| | ||: | |: : |:: : |::

Db 642 VYCFVESLLLCRAVMLASIGSEGGFLSHVN---GYIQAVICAFSVVAQVGHVSQQLLVVM 697

QY I 3 LV RGS GEG F QGN SSG WAW ASP CAT AN M D F - - - - - V M A L I I V M L L L L G A F L G A W P A L C Z S

Db 698 HV-----ASETVCSENIYYGRNWLGLLAYDFALLCCVGALI---PSIY 737

Qy 233 GRYKRWKHGVFLTTATATSAIWWVWVIMVITYGKQKHNSTWDDPTLAIALAANAWAFV 292

Db 738 RSQRNYRE-GILIVIGSVLIMVIWVAWIALSLFGE-----WRDAAIPLGLQASGWAVL 790

QY 293 LFYVIP-----EVSQVTKSSPEQS-----YQGDMPYTRGVGYETI 327

Db 791 VGIIPRTFLIVRGIERSDIAQALPULTSLAFQNNQYSSESQSVYECV 838

RESULT 15

Q9V4U3
ID Q9V4U3 PRELIMINARY; PRT; 738 AA.

AC Q9V4U3;
DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)

DE CG8692 PROTEIN.
GN CG8692.

OS *Drosophila melanogaster* (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.

OX	NCBI_TaxID=7227;
RN	[1]

RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;

RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.

RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshreti A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,

```
OX NCBI_TaxID=10090;
RN SEQUENCE FROM N.A.
RP "Metabotropic glutamate-like sequence in C. elegans.";
RC TISSUE=BRIN;
RA MEDLINE=20145471; PubMed=10679242;
RX Inoue S., Sano H., Ohka M.;
RT "Growth suppression of Escherichia coli by induction of expression of
RL mammalian genes with transmembrane or ATPase domains.";
RR Biochem. Biophys. Res. Commun. 268:553-561(2000).
RD EMBL; AB030197; BAA92760.1; -.
RE MGD; MGI:1927596; Gprc5b.
RF NON_TER
FT
SQ SEQUENCE 262 AA; 29305 MW; 6407A0D9FADF75F8 CRC64;

Query Match 18.3%; Score 425.5; DB 11; Length 262;
Best Local Similarity 38.5%; Pred. No. 1.2e-27;
Matches 102; Conservative 38; Mismatches 72; Indels 53; Gaps 8;

Qy 144 LARKNHGPRGWVFTVALLTLVEVIINTEWLIITLVRSGEQGGPQGNSSAGWAVASPCA 203
Db 2 LVROGTSPASQVSLALCLMLQVVIITATEWLTLVLRDT-----KPCA 46

Qy 204 IANMDFVALLIYVMLLLGALGAWPALCGRYKRWKRGVFLVLTATTSVAIVVWVIMY 263
Db 47 YEPDMFVALLIYVMLLLGALGAWPALCGRYKRWKRGVFLVLTATTSVAIVVWVIMY 106

Qy 264 TYGN--KOHNSPTWDDPTLAIALAANAWAFVLFVIVPEVSQVTKSSPEQSQGDMYPT- 319
Db 107 LFGSLIKQGD--WSDPTLAITLAASGWVFIHAIPEI-HYTLPLQENPNFYDTS 163

Qy 320 ----RGVGYETILKEQSQSMFVENKAFSMDPEVAAKRPVSPYSGYNGQL----- 365
Db 164 QPRMRETAFD--EEMHLPAYMENKAFSMDPEVAAKRPVSPYSGYNGQL----- 218

Qy 366 -----LTSVQPTMALM 378
Db 219 PSSLGRLPSAPFRSNVQPTMAV 243

RESULT 12
Q20073 PRELIMINARY; PRT; 1404 AA.
AC Q20073; 001363;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 157.9 KDA PROTEIN (CELF35-1).
GN F35H10.10.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Woessner J.;
RT "The sequence of C. elegans cosmid F35H10.";
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
```

```
RN SEQUENCE FROM N.A.
RP Abe T., Tanemoto M., Nishio T., Hebert S.C.;
RT "Metabotropic glutamate-like sequence in C. elegans.";
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U00934; AAA81683.2; -.
DR EMBL; U00870; AAB51198.1; -.
DR InterPro; IPR000337; GPCR_Mgr.
DR InterPro; IPR000847; HTH_LysR.
DR Pfam; PF00003; 7tm_3; 1.
DR PRINTS; PR00248; GPCRMR.
DR PROSITE; PS0259; G_PROTEIN_RECEP_F3_4; 1.
DR PROSITE; PS00044; HTH_LYSR_FAMILY; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 1404 AA; 157943 MW; 46D8282F7FC464A0 CRC64;

Query Match 7.2%; Score 167; DB 5; Length 1404;
Best Local Similarity 20.8%; Pred. No. 3.5e-05;
Matches 70; Conservative 55; Mismatches 140; Indels 72; Gaps 11;

Qy 46 SGAWG-----IVLEAVAGAGIVTFTVLTITLVASLFPVODTKKRSLLTQVFPFLGLTGL 99
Db 1015 TGHWRDHPHNVLVLLALITLVVVAIAVLVLVVKL--YLRVVKGNQSLG--ISLLIGIII 1070

Qy 100 LFCLVFACVVKPDFSTCASRRFLFGLFAICFSCLAARHVFALNFKRNHGRGWVFTV 159
Db 1071 LYSTAFFEVDPTDSVCRRLRVILHGLGVTICFVGMIAKATOLRNAETLGF----- 1121

Qy 160 ALLTLVEVIINTEWLIITLVRSGEQGGPQGNSSAGWAV-----ASPCA 204
Db 1122 ----TAIHSEFWNYWLLLFET-----VGQIALSISFLEPFMTSTIGVIDTNVQRMCTM 1172

Qy 205 ANMDFVALLIYVMLLLGALGAWPALCGR-YKRWKRGVFLVLTATTSVAIVVWVIMY 263
Db 1173 GKVEFVSNFYVMIIL---FMALFISMLNRNKRNYKTKWLLYSTVLCFTWVAVITLY 1229

Qy 264 TYGNKQNSPTWDDPTLAIALAANAWAFVLFVIVPEVSQVTKSSP-EQSQGDMYPTRGV 322
Db 1230 LVLDHE-----FRDTVIVVELVACATILGLFGPKVILLSEYEPVVVAFKRDPEPN--- 1281

Qy 323 GYETILKEQSQSMFVENKAFSMDPEVAAKRPVSPYS 359
Db 1282 -----HTDLFEKDDLPQRAVSPAS 1302

RESULT 13
Q20073 PRELIMINARY; PRT; 879 AA.
AC Q20073;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE METABOTROPIC GLUTAMATE RECEPTOR 3 PROTEIN.
GN GRM3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV;
RX MEDLINE=20012997; PubMed=10544282;
RA Minoshima T., Nakanishi S.;
RT "Structural organization of the mouse metabotropic glutamate receptor
subtype 3 and its regulation by growth factors in cultured cortical
astrocytes.";
RL J. Biochem. 126:889-896(1999).
DR EMBL; AF170701; AAF06741.1; -.
DR EMBL; AF170697; AAF06741.1; JOINED.
DR EMBL; AF170698; AAF06741.1; JOINED.
DR EMBL; AF170699; AAF06741.1; JOINED.
DR EMBL; AF170700; AAF06741.1; JOINED.
```

DR EMBL: AF209923; AAF72873.1; -.
DR InterPro: IPR002956; Bridge_of_7less.
DR InterPro: IPR000337; GPCR_Mgr.
DR Pfam: PF00003; 7tm_3; 1.
DR PRINTS: PR01223; BRIDBOF7LESS.
KW Receptor.
SQ SEQUENCE 345 AA; 38791 MW; 14B09E4C5E9F9B49 CRC64;

Query Match 22.4%; Score 521.5; DB 4; Length 345;
Best Local Similarity 33.2%; Pred. No. 1.6e-35;
Matches 125; Conservative 64; Mismatches 124; Indels 63; Gaps 11;

QY 39 YNLCDSRGAWGIVLEAVAGAGIVTFTVLTILVASLFFVQDTKKRSLLGTQVFFLLGTL 98
Db 12 YFLCDAGPGWIIIESLAILGIVVTILLALLAFLLMRKIQDCSQMNVLPQTQLFLLSVL 71

QY 99 GLFCLVFACVVKPDFSTCASRRFLFGLVFAICFSLAAHVAFALNFKNKGPRGWIFT 158
Db 72 GLGLAFAPAFIELNQQTAPVRYFLFGLVFAICFSLAAHASNLVKLV-----RGCVSFS 125

QY 159 -----VALLLTVEVIINTEWLIITLVRSGGGPOGNSAGWAVASPCAIAANMDFVMA 212
Db 126 WTTILCIAIGCSLLQIIATEYVTLITRG-----MMFVNMTPCQL-NVDEWVL 173

QY 213 LIYVMLLLGAFGLGAWPALCGRYKRWKRGVFLVLTATSVAIWVWVIMTYGKQ-HN 271
Db 174 LVYVFLMALTFVSKATFCGPCENKQHGRLIFATVLSIIWVWISMLLRGNPQFOR 233

QY 272 SPTWDDPTALALANAWAFVLEFVPIEVSQVTKSSPEQ-SYQGDMPYTRGVETILKE 330
Db 234 QPWDDPVCIALTNWVFLLYIVPELCILYRSCQECPLQGNACPV--TAYQ----- 286

QY 331 QKGOSMEVKNKAFS-----MDEPVAAKRPVSPYSYQGLTTSVYQPTMALMHKVP 383
Db 287 ---HSFQVENQELSRSDSGAEDVA-----LTSYGTPIQPTQVDPDQ----- 327

QY 384 EGAYDIILPRATANSQ 399
Db 328 -----ECFIPQAKLSQ 339

RESULT 9
Q9JIL6 PRELIMINARY; PRT; 300 AA.

AC Q9JIL6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE ORPHAN G-PROTEIN COUPLED RECEPTOR.
GN GPRC5D.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21210966; PubMed=11311935;
RA Brauner-Osborne H., Jensen A.A., Sheppard P.O., Brodin B.,
RA Krogsgaard-Larsen P., O'Hara P.;
RT "Cloning and characterization of a human orphan family C G-protein
RT coupled receptor GPRC5D(1).";
RL Blochim. Biophys. Acta 1518:237-248(2001).
DR EMBL: AF218809; AAF72880.1; -.
DR MGD; MGI:1935037; Gprc5d.
DR InterPro: IPR002956; Bridge_of_7less.
DR DR InterPro: IPR000337; GPCR_Mgr.
DR Pfam: PF00003; 7tm_3; 1.
DR PRINTS; PR01223; BRIDBOF7LESS.
KW Receptor.
SQ SEQUENCE 300 AA; 34488 MW; D9423064260A4C59 CRC64;

Query Match 21.0%; Score 487.5; DB 11; Length 300;
Best Local Similarity 38.4%; Pred. No. 9.3e-33;
Matches 109; Conservative 44; Mismatches 112; Indels 19; Gaps 4;

QY 39 YNLCDSRGAWGIVLEAVAGAGIVTFTVLTILVASLFFVQDTKKRSLLGTQVFFLLGTL 98
Db 12 YFLCDFNEGPAIIVLESIAVIGIVVTILLALLAFLLMRKYQDCSQMNVLPQTQLFLLA 71

QY 99 GLFCLVFACVVKPDFSTCASRRFLFGLVFAICFSLAAHVAFALNFKNKGPRGWIFT 158
Db 72 GLGLTFAFIQLNQTAPVRYFLFGLVFAICFSLAAHASNLVKLVGRVRSFCWTTILF 131

QY 159 VALLLTVEVIINTEWLIITLVRSGGGPOGNSAGWAVASPCAIAANMDFVMAIYVML 218
Db 132 IAGVSLQTIILAEYVTLITRG-----LMFEHMTPIQL-NVDEVCULLIVLF 179

QY 219 LLLGAFGLGAWPALCGRYKRWKRGVFLVLTATSVAIWVWVIMTYGKQ-HNSPTWDD 277
Db 180 LMALTFVSKATFCGPCENKQHGRLIFATVLSIIWVWISMLLRGNPQORQPHWDD 239

QY 278 PTLAIALANAWAFVLEFVPIEVSQVTKSSPEQSYQGDMPYTRG 321
Db 240 AVICIGLVTNAWVFLIYIPELSILYRSCQEC-----PTQG 277

RESULT 10
Q9BSF0 PRELIMINARY; PRT; 105 AA.

AC Q9BSF0;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE SIMILAR TO G PROTEIN-COUPLED RECEPTOR, FAMILY C, GROUP 5, MEMBER C.
DE C.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PANCREATIC ADENOCARCINOMA;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC004925; AAH04925.1; -.
KW Receptor.
SQ SEQUENCE 105 AA; 11107 MW; D8B3FA9FA2559C99 CRC64;

Query Match 20.6%; Score 478; DB 4; Length 105;
Best Local Similarity 30.3%; Pred. No. 1.6e-32;
Matches 93; Conservative 2; Mismatches 4; Indels 4; Gaps 1;

QY 1 MAIHKALVMCLGLPLFLFPGAWAGHVPYPCGSQLNPLYYNLCDSRGAWGIVLEAVAG 60
Db 1 MAIHKALVMCLGLPLFLFPGAWAGHVPYPCGSQLNPLYYNLCDSRGAWGIVLEAVAG 60

QY 61 IVTTFVLTILVASLFPVQDTKKRSLLGTQV-----FFLLGTGL 99
Db 61 IVTTFVLTILVASLFPVQDTKKRSLLGTQLRGRCHHTAGTMG 103

RESULT 11
Q9JMF0 PRELIMINARY; PRT; 262 AA.

AC Q9JMF0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE UNKNOWN PROTEIN (FRAGMENT).
GN GPRC5B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.


```

Db 118 RFLWGVLFAFCSCLLSQAWVRRLRHGTGPAGWQLVGLALCLMLVQVITIAVEWLVLT 177
      |||:|||||:|||||: : : || || || : || || || || || || || || || || ||
QY 180 VRSGEGGPGNSSAGWAVASPCAIANMDFVMAIIYVMLLLGLGAFGLAWPALCGRYKRW 239
      |||:|||||:|||||: : : || || || : || || || || || || || || || || ||
Db 178 LRDT-----RPAAYEPDMFVMAIIYDMLLVVTVTLGALFTLCGKFKRWK 222
      |||:|||||:|||||: : : || || || : || || || || || || || || || || ||
QY 240 KHGVFLLTATSVAIWVIMVITYGN-KOHNSPTWDDPTLAIALAANAWAFVLFVVIP 298
      |||:|||||:|||||: : : || || || : || || || || || || || || || || ||
Db 223 LNAFLITAFSLVLIWAWMTYLFNGVKLQOGDANDPTLAIITLAASGWVFIHAIP 282
      |||:|||||:|||||: : : || || || : || || || || || || || || || || ||
QY 299 EV-----SOVTKSSP---EQSYQGDMPYTRGVGYETILKEQKQSMFVENKA 342
      |||:|||||:|||||: : : || || || : || || || || || || || || || || ||
Db 283 EIHTLLPALQENTPNYFDTSPMRMETAPEEDVQLPRA-----YMENKA 327
      |||:|||||:|||||: : : || || || : || || || || || || || || || || ||
QY 343 FSDMEPVAA-----KRPVSPSYSGYNGQLLTSVYQPTMALM 378
      |||:|||||:|||||: : : || || || : || || || || || || || || || || ||
Db 328 FSDMEHNAALRTAGFPNGSLGKRPSGLGRPSAPER-----SNVYQPTMAV 376
      |||:|||||:|||||: : : || || || : || || || || || || || || || || ||

RESULT 4
ID 075205 PRELIMINARY; PRT; 403 AA.
AC 075205;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE HYPOTHETICAL 44.8 KDA PROTEIN (GPCR5B PROTEIN).
GN A-69G12.1 OR GPCR5B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RA Adams M.D., Loftus B.J., Zhou L., Crosby M., Fuhrmann J., Mason T.M.,
RA Brandon R., Kim U.J., Kerlavage A.R., Venter J.C.;
RA "Homo sapiens Chromosome 16 BAC clone C1987SK-A-69G12.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RA Wei H., Osborne B., Spruyt M., Murphy D.;
RC TISSUE=BRAIN;
RT "Cloning of a novel G protein-coupled receptor localized on human
RT chromosome 16p12.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE FROM N.A.
RA Robbins M.J., Michalovich D., Hill J., Calver A.R., Medhurst A.D.,
RA Gloger I., Sims M.A., Middlemiss D.N., Pangalos M.N.;
RT "Molecular cloning and characterisation of two novel retinoic acid
RT inducible orphan G-protein coupled receptors.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC004131; AAC27544.1; -.
DR EMBL; AF181862; AAF05331.1; -.
DR EMBL; AJ276101; CAC00632.1; -.
DR InterPro; IPR000337; GPCR_Mgr.
DR Pfam; PF00003; 7tm3; 1.
DR PROSITE; P550259; G_PROTEIN_RECIP_F3_4; 1.
KW Hypothetical protein; Receptor.
SQ SEQUENCE 403 AA; 44795 MW; 3902A16C4F69C26E CRC64;

```

```

Query Match 31.5%; Score 733; DB 4; Length 403;
Best Local Similarity 39.4%; Pred. No. 4.4e-53;
Matches 164; Conservative 60; Mismatches 114; Indels 78; Gaps 9;

QY 1 MATHKALVCLGLPLFLFPG-AWAQGHVPPGCSQGLNPLYNLCDSGAWGVLEAVAGA 59
      |||:|||||:|||||: : : || || || : || || || || || || || || || || ||
Db 9 MRAHQVLTFL---LPIVTSVASENASTSRGCGLDLLPQVYVSLCDLDAIWGIVVEAVAGA 65
      |||:|||||:|||||: : : || || || : || || || || || || || || || || ||
QY 60 GIYTFVLITILVASLPFVQDTKKRSLGTQVFFLLGLTGLFLCVACVKKPDEFSTCASR 119
      |||:|||||:|||||: : : || || || : || || || || || || || || || || ||

```

```

Db 66 GALITLLMLLLVRLPFIKEKSKSPVGLHFLFLGLTGLGFLGTFAFIIQEDETICSVR 125
QY 120 RELFGVLFAFCSCLAHVFALNFLARKNHGPRGWIFTVALLLTVEVINTEWLITL 179
      |||:|||||:|||||: : : || || || : || || || || || || || || || || ||
Db 126 RFLWGVLFALFCSCLLSQAWVRRLRHGTGPAGWQLVGLALCLMLVQVITIAVEWLVLT 185
      |||:|||||:|||||: : : || || || : || || || || || || || || || || ||
QY 180 VRSGEGGPGNSSAGWAVASPCAIANMDFVMAIIYVMLLLGLGAFGLAWPALCGRYKRW 239
      |||:|||||:|||||: : : || || || : || || || || || || || || || || ||
Db 186 LRDT-----RPAAYEPDMFVMAIIYDMLLVVTVTLGALFTLCGKFKRWK 230
      |||:|||||:|||||: : : || || || : || || || || || || || || || || ||
QY 240 KHGVFLLTATSVAIWVIMVITYGN-KOHNSPTWDDPTLAIALAANAWAFVLFVVIP 298
      |||:|||||:|||||: : : || || || : || || || || || || || || || || ||
Db 231 LNAFLITAFSLVLIWAWMTYLFNGVKLQOGDANDPTLAIITLAASGWVFIHAIP 290
      |||:|||||:|||||: : : || || || : || || || || || || || || || || ||
QY 299 EV-----SOVTKSSP---EQSYQGDMPYTRGVGYETILKEQKQSMFVENKA 342
      |||:|||||:|||||: : : || || || : || || || || || || || || || || ||
Db 291 EIHTLLPALQENTPNYFDTSPMRMETAPEEDVQLPRA-----YMENKA 335
      |||:|||||:|||||: : : || || || : || || || || || || || || || || ||
QY 343 FSDMEPVAA-----KRPVSPSYSGYNGQLLTSVYQPTMALM 378
      |||:|||||:|||||: : : || || || : || || || || || || || || || || ||
Db 336 FSDMEHNAALRTAGFPNGSLGKRPSGLGRPSAPER-----SNVYQPTMAV 384
      |||:|||||:|||||: : : || || || : || || || || || || || || || || ||

RESULT 5
Q92320 PRELIMINARY; PRT; 410 AA.
AC Q92320;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ORPHAN G PROTEIN-COUPLED RECEPTOR.
GN RAIG2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN SEQUENCE FROM N.A.
RA Tao Q., Lotan R.;
RT "Molecular cloning and characterization of mouse retinoic acid-
RT inducible orphan G protein-coupled receptors.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF378831; AAK58076.1; -.
KW Receptor.
SQ SEQUENCE 410 AA; 45899 MW; E1DA5283270FF34 CRC64;

```

```

Query Match 30.8%; Score 716.5; DB 11; Length 410;
Best Local Similarity 40.7%; Pred. No. 1.1e-51;
Matches 162; Conservative 62; Mismatches 119; Indels 55; Gaps 9;

QY 13 LPFLFPGAWAQAQGHVPP--GCSQGLNPLYNLCDSGAWGVLEAVAGAGIVTFVLTII 70
      |||:|||||:|||||: : : || || || : || || || || || || || || || || ||
Db 17 LPLLLVIASVASENASTSRGCGLDLLPQVYVSLCDLDAIWGIVVEAVAGAGALITLLMLJ 76
      |||:|||||:|||||: : : || || || : || || || || || || || || || || ||
QY 71 LVASLPVQDTKKRSLGTQVFFLLGLTGLFLCVACVKKPDEFSTCASRFLFGVLPFAIC 130
      |||:|||||:|||||: : : || || || : || || || || || || || || || || ||
Db 77 LLVRLPFIKDKERRKRPVCLHFLGLTGLGFLGTFAFIIQEDETICSIRRLGVLGVLFALC 136
      |||:|||||:|||||: : : || || || : || || || || || || || || || || ||
QY 131 FSLCLAAHVFALNFLARKNHGPRGWIFTVALLLTVEVINTEWLITLVRSGEGGPGQ 190
      |||:|||||:|||||: : : || || || : || || || || || || || || || || ||
Db 137 FSCLLSQAWVRRLVRQGTSPASQQLVSLALCLMLVQVITIAVEWLVLTLDRT----- 189
      |||:|||||:|||||: : : || || || : || || || || || || || || || || ||
QY 191 NSSAGWAVASPCAIANMDFVMAIIYVMLLLGLGAFGLAWPALCGRYKRWKGVFLTTA 250
      |||:|||||:|||||: : : || || || : || || || || || || || || || || ||
Db 190 -----KPAAYEPDMFVMAIIYDMLLVVTVTLGALFTLCGKFKRWKNGAFILVTF 241
      |||:|||||:|||||: : : || || || : || || || || || || || || || || ||
QY 251 TSAIVWVIMVITYGN---KOHNSPTWDDPTLAIALAANAWAFVLFVLPVSQVTKSS 307
      |||:|||||:|||||: : : || || || : || || || || || || || || || || ||
Db 242 LSAIIVWVIMVITYLFGNSLIKQGDA--WSDPTLAIITLAASGWVFIHAIPETI-HYTL 298
      |||:|||||:|||||: : : || || || : || || || || || || || || || || ||
QY 308 PEQSYQGDMPY-----RGVGYETILKEQKQSMFVENKAFSMDPEPVAARKRPVSPSY 362
      |||:|||||:|||||: : : || || || : || || || || || || || || || || ||

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2002, 09:32:51 ; Search time 31.28 seconds
(without alignments)
2438.963 Million cell updates/sec

Title: US-09-895-686-1

Perfect score: 2326

Sequence: 1 MAHKALVMCLGLPLFLPG.....ATPPKDGKNSQVFRNPYWD 441

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriaph:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2326	100.0	441	4 Q9NZG5	Q9nzg5 homo sapien
2	2325	100.0	441	4 Q9NQ84	Q9nq84 homo sapien
3	733	31.5	395	4 Q9NZH0	Q9nzh0 homo sapien
4	733	31.5	403	4 Q75205	Q75205 homo sapien
5	716.5	30.8	410	11 Q92320	Q92320 mus musculus
6	651	28.0	162	4 Q9NXI0	Q9nxi0 homo sapien
7	558	24.0	357	4 Q95357	Q95357 homo sapien
8	521.5	22.4	345	4 Q9NZD1	Q9nzd1 homo sapien
9	487.5	21.0	300	11 Q9JIL6	Q9jil6 mus musculus
10	478	20.6	105	4 Q9BSP0	Q9bsp0 homo sapien
11	425.5	18.3	262	11 Q9JMF0	Q9jmf0 mus musculus
12	167	7.2	1404	5 Q20073	Q20073 caenorhabditis
13	160.5	6.9	879	11 Q9QYS2	Q9qys2 mus musculus
14	155	6.7	917	5 Q24265	Q24265 drosophila
15	154	6.6	738	5 Q9V4U3	Q9v4u3 drosophila
16	151	6.5	872	4 Q9H3N6	Q9h3n6 homo sapien

17	150.5	6.5	877	4 Q9UGT0	Q9ugt0 homo sapien
18	143	6.1	1199	11 Q9EPV6	Q9epv6 mus musculus
19	140	6.0	589	4 Q9UGS9	Q9ugs9 homo sapien
20	133	5.7	1218	13 Q90ZF3	Q90zf3 oncorhynchus
21	132.5	5.7	551	13 Q9PWQ0	Q9pwq0 fugu rubrip
22	132	5.7	251	13 Q73657	Q73657 fugu rubrip
23	129.5	5.6	251	13 Q73656	Q73656 fugu rubrip
24	129	5.5	583	5 Q9VNZ5	Q9vnz5 drosophila
25	128.5	5.5	877	13 Q9PW88	Q9pw88 carassius a
26	123	5.3	875	13 Q73640	Q73640 fugu rubrip
27	121.5	5.2	251	13 Q73653	Q73653 fugu rubrip
28	119.5	5.1	251	13 Q73654	Q73654 fugu rubrip
29	119.5	5.1	251	13 Q73655	Q73655 fugu rubrip
30	119	5.1	250	13 Q73646	Q73646 fugu rubrip
31	117.5	5.1	699	16 Q9ABL0	Q9abl0 caulobacter
32	117	5.0	864	13 Q73637	Q73637 fugu rubrip
33	116	5.0	940	13 Q73635	Q73635 fugu rubrip
34	115.5	5.0	2426	3 Q96U00	Q96u00 aspergillus
35	115	4.9	458	13 Q93555	Q93555 carassius a
36	115	4.9	856	13 Q73638	Q73638 fugu rubrip
37	115	4.9	983	11 Q62916	Q62916 rattus norv
38	114	4.9	844	13 Q93552	Q93552 carassius a
39	113	4.9	250	13 Q73647	Q73647 fugu rubrip
40	113	4.9	848	13 Q93553	Q93553 carassius a
41	112.5	4.8	362	13 Q9PSY1	Q9psy1 carassius a
42	112.5	4.8	408	13 Q93558	Q93558 carassius a
43	112	4.8	390	2 Q9L949	Q9l949 pseudomonas
44	110.5	4.8	250	13 Q73650	Q73650 fugu rubrip
45	110	4.7	250	13 Q73651	Q73651 fugu rubrip

ALIGNMENTS

RESULT	1
Q9NZG5	
ID	Q9NZG5 PRELIMINARY; PRT; 441 AA.
AC	Q9NZG5;
DT	01-OCT-2000 (TREMBLrel. 15, Created)
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE	ORPHAN G-PROTEIN COUPLED RECEPTOR (G PROTEIN-COUPLED RECEPTOR, FAMILY
DE	C. GROUP 5, MEMBER C).
GN	GPRC5C.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=21210966; PubMed=11311935;
RA	Brauner-Osborne H., Jensen A.A., Sheppard P.O., Brodin B.,
RA	Krogsgaard-Larsen P., O'Hara P.;
RT	"Cloning and characterization of a human orphan family C G-protein
RT	coupled receptor GPRC5D(1).";
RL	Biochim. Biophys. Acta 1518:237-248(2001).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	TISSUE=PANCREAS, EPITHELIOID CARCINOMA;
RA	Strausberg R.;
RL	Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AF207989; AAF72870.1; -.
DR	EMBL; BC016860; AAH16860.1; -.
DR	InterPro; IPR003037; GPCR_Mgr.
DR	InterPro; IPR003006; Ig_MHC.
DR	Pfam; PF00003; 7tm3.1.
DR	PROSITE; PS50259; G-PROTEIN_RECEP_F3_4; 1.
DR	PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW	Receptor.
SQ	SEQUENCE 441 AA; 48193 MW; B255D9680328FE3D CRC64;
Query Match	100.0%; Score 2326; DB 4; Length 441;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2002, 14:34:17 : Search time 1947.48 Seconds
(without alignments)
12606.328 Million cell updates/sec

Title: US-09-895-686-7

Sequence: 1819
1 cggctcggcgcctccaccagc.....ctttatctactcttaaaaa 1819

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 674847542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estda:*
2: em_esthum:*
3: em_estlin:*
4: em_estlmu:*
5: em_estlov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gD_estl:*
10: gD_est2:*
11: gD_hlc:*
12: gD_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pla:*
16: em_gss_vtl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1081	59.4	1815	11	AK014308 Mus muscu
2	885.6	48.7	903	9	AL545009 AL545009
3	882.8	48.5	955	9	AL570913 AL570913
4	852.6	46.9	906	9	AL577178 AL577178
5	847	46.6	874	9	AL556748 AL556748
6	845.2	46.5	905	10	BI916185 BI916185
7	837.6	46.0	851	9	AL546406 AL546406
8	792.6	43.6	842	9	AL571754 AL571754
9	775.2	42.6	956	10	BG829477 BG829477
10	761.2	41.8	846	9	AL578937 AL578937
11	747.8	41.1	873	10	BI769602 BI769602
12	744.4	40.9	820	10	BI834606 BI834606
13	732	40.2	933	10	BI772730 BI772730
14	728.4	40.0	776	10	BG750651 BG750651
15	710.2	39.0	1052	10	BI822698 BI822698
16	687	37.8	689	10	BI764504 BI764504
17	678.8	37.3	710	9	AM954136 AM954136

18	607.2	33.4	685	10	BE898354 BE898354
19	588.8	32.4	623	9	AA877534 AA877534
20	558.4	30.7	698	10	BI760442 BI760442
21	529	29.1	653	10	BI400350 BI400350
22	517.2	28.4	849	10	BE909608 BE909608
23	479.2	26.3	580	10	BI404524 BI404524
24	478.6	26.3	575	10	BE683495 BE683495
25	475	26.1	478	9	AM959466 AM959466
26	472.4	26.0	588	10	BG692011 BG692011
27	459	25.2	960	9	AI921413 AI921413
28	458.2	25.2	400	10	BF786181 BF786181
29	456.6	25.1	959	10	BF783102 BF783102
30	455.8	25.1	459	9	AI075186 AI075186
31	451.6	24.8	555	10	BI402962 BI402962
32	443	24.4	443	9	AI298839 AI298839
33	441.6	24.3	548	9	AA178961 AA178961
34	436	24.0	436	9	AI659014 AI659014
35	431.8	23.7	443	9	AI016010 AI016010
36	427.2	23.5	444	9	AI694173 AI694173
37	427.2	23.5	513	10	BM107227 BM107227
38	423.2	23.3	566	9	BE232134 BE232134
39	422.4	23.2	432	9	AI052635 AI052635
40	422	23.0	510	10	BI403847 BI403847
41	419.2	22.9	550	10	BG082452 BG082452
42	416.6	22.9	522	10	BI338854 BI338854
43	415	22.8	503	10	BG610066 BG610066
44	408.6	22.5	606	10	BI825005 BI825005
45	405.4	22.3	442	9	AA902120 AA902120

ALIGNMENTS

RESULT 1	AK014308	1815 bp	mus musculus 14, 17 days embryo head cDNA, enriched library, clone:320002M13:homolog to ORFHAN G-PROTEIN
LOCUS	AK014308		
DEFINITION	AK014308	1815 bp	mus musculus 14, 17 days embryo head cDNA, enriched library, clone:320002M13:homolog to ORFHAN G-PROTEIN
ACCESSION	AK014308	GI:12852069	
VERSION	AK014308.1		
KEYWORDS	HTC: CAP trapper.		
SOURCE	Mus musculus (strain:C57BL/6J) 14, 17 days embryo head cDNA to mRNA, clone.lib:RIKEN full-length enriched mouse cDNA library		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	Carninci,P. and Hayashizaki,Y.		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)		
MEDLINE	99279253		
PUBMED	10349636		
REFERENCE	2 (sites)		
AUTHORS	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.		
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		
MEDLINE	20499374		
PUBMED	11042159		
REFERENCE	3 (sites)		
AUTHORS	Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuami,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsunoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujikake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Mataliki,M., Onoda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsunaga,S., Kawai,D., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.		
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipillar sequencer		
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)		

MEDLINE 20530913
 PUBMED 11076861
 REFERENCE 4 (sites)
 AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
 TITLE Functional annotation of a full-length mouse cDNA collection
 JOURNAL Nature 409, 685-690 (2001)
 REFERENCE 5 (bases 1 to 1815)
 AUTHORS Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,

Aikawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bull, C., Caranluci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hirooka, T., Horii, F., Hume, D., Imocanci, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Schemm, L., Shibata, K., Shibata, Y., Shilagawa, A., Shiraki, T., Sobabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

TITLE Direct Submission
JOURNAL Submitted (10-Jun-2000) Yoshinide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URI: <http://genome-gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT Please visit our web site (<http://genome-gsc.riken.go.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGGAGACGAGAATCCAAAGCGCTCTTTTCTTTTTTTCVN 3'], cDNA was prepared by using reverse transcriptase thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGGAGACGATTCTCAGACTTAATTAATTAATCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOLR.

```

FEATURES
source
location/qualifiers
1..1815
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="MGD:MGI:1902714"
/db_xref="Enxox:10090"
/clone="320002M13"
/tissue_type="head"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="14, 17 days embryo"
1..1815
misc_feature
/note="data source:SPTR, source key:Q9NZC5, evidence:ISS
homolog to ORPHAN G-PROTEIN COUPLED RECEPTOR"
BASE COUNT
342 a
544 c
503 g
426 t
ORIGIN

```

Query Match	59.4%	Score 1081;	DB 11;	Length 1815;
Best Local Similarity	79.4%	Pred. No. 1.1e-218;		
Matches 1415: Conservative	0;	Mismatches 335;	Indels 33;	Gaps 10;

QY	48	agggaccaccaacagagcctgagcctgagagccagaatgagccatccacaagaacctgtgcat	107
Db	42	AGGCACCCACGCTGAAGCCTGGCTCTGGGAGACCGAGATGGCACTCACGAACCTTGGCTGAT	101
QY	108	gtgcctggaactgactctctctccctggttcaccaagagccttggcccaagagccatgtgccacc	157
Db	102	GTGGCTTGAAGAACTGGCTCTCTTC---TTCCACGAGAGCCCTTGGGCCAGAAATATGGCCACC	158

QY	168	cggctgcagccaagcgctcaacccctgtatctacaacctgtgtgaecgcctctgggcgctg	227
Db	159	TGGTGGACGCCCAATATGTGATCCCCCTCTACTCAACCTCTGTGGACCGCTGGGCCCTG	218
QY	228	gggcaatcctcttgaaagccgttgagcttgaggcggaatgttacaagaagttgtgtcacaat	287
Db	219	GGGCATTTGCTTTGGAGGACGTGGCTGGAGGACAGGCATCATCAGCACATTGTGTAAACAT	278
QY	288	caccctgtgtgcgaagctccccccttgtgcagagacacccaagaaacggagccctgtggagac	347
Db	279	CATCCTTGTGGCTAGCCCTTCACATTTGTGGACGAGACATAAGAAACGGGA - CCTCTTGGGGAC	337
QY	348	ccaagtatctctctctctcttgaggaaacctgggacctctctgcctcgtytttgccctgtgtgt	407
Db	338	CCAGGTGTTCTTCCGTGGGGACCCCGGGGCTCTCTCTGCTGCTGTGTTGGCTGTGTGT	397
QY	408	gaagccgaactctccacccctgtgacctctggagacctccctcttgagggtctcttgccacat	467
Db	398	GAAGCCGGACTTCTCTACCTGTGGCTTCCACCGCTTTCCTTTGGGGTCTCTGTTTGCCAT	457
QY	468	ctgctctctctgtgtggccgagctcaagctctcttgccctcaactctctgtggccggaaagaca	527
Db	458	CTGGCTTCCTGTTGGTAGTGACGTCCTTTCCTTCACCTTCCTTAACCCGGAGAACCA	517
QY	528	cgggcccgggagctgggtgatcttcaactgtgtgcctgtgcctgtgcagaccgtgtagaagtgat	587
Db	518	TGGGGCCCCGAGGCGTGGGTGATCTTTCACCCGTGGGGCTGTGCTCACCTGTGTGAGTCAAT	577
QY	588	catacaatacaagatggctgtgatcacaccctggctctggaggagatgggagagagcgccctca	647
Db	578	CATTACACCGAGTGGCTTATCATACCTCTGGTACGGGAGGTGGCCAGGTTAAGCCCTT	637
QY	648	gggcacacagcagcgcaagcgctggggccgtgtgcctcccccctgtgcacatcgccaacatlggact	707
Db	638	GGGCAATGTACAGTCCCGACCTCGACATGACATGACCTCTCCGTGGGCATGGCCAAATGAGACTT	697
QY	708	tgatcatggcaatatctcaagtcgatgctgcctgcctgcctgggtgtgcctctctctggggcctggcc	767
Db	698	TGTCTATGGCTCTCTCATCTACCTAAAGCTGTTGGTGGCTGACGGCCCTTCTAAGGAGCCTGGGC	757
QY	768	cgccctgtgtggccgctcaacaagcgctggcctaaagcatlggggtctcttgtctctcaacac	827
Db	758	CACCTTGTGTGGCCGCTTCAAGCCGCTTGGCGGAACAGGGGCTTGTGTGTGCTACACAC	817
QY	828	agccaacctcgttgtgcatalatgggtgtgtgtgagatcgctcatgtatctacacggaacaagca	887
Db	818	TGTCTATCTTCATCCGCACTGTGGGTGATGATGTTCATGTATACCTACGCGACGACGAGA	877
QY	888	gcaacaacgtctcccaacctgggagatgaacccaagctgtggcaatcgcccttcgcccgaatgactgt	947
Db	878	GCACCATATGGCCCCACCTGGGATGACCCCAAGTTGGGCATTTGCCCTGTGCGCAATGGCTGG	937
QY	948	ggacctgtctctctctctcaagtcaccccccgaggtctcccaagtgbaaccaagtctcagccca	1007
Db	938	GACCTTTGTCTCTTCTTATGTGCAAT - CTTGAGGCTCTACAGGTBACCMAACTTAAGCCAGA	996
QY	1008	gcaaaagctacccaagggagacatgtaccccaaccccgggcgctggggctatagacaatccctgaa	1067
Db	997	ACAGAGCTACCA - GGGGACATGTATACCGGACCGAGGGGTGGGCTATGAGACATCTGAA	1055
QY	1068	agagcagagaaggtcagaagcatgtctgtgagaaacaagccctttcocalatgatagtgcggt	1127
Db	1056	GGACAGACGGGGCCAGAGCATGTTTGTGGAAACAAGGCAATTTTCATATGAGCAACAGAC	1115
QY	1128	tgaagcttaagagggccggtgtgcacacataagccggtacaaatgggcaagctgtctgaacaaggt	1187
Db	1116	CTTCAGCAAAAGACACTGTTCACCTTAACATGTGCTCAATATGACACAGCTCTATACCAGCT	1175
QY	1188	gtaccagcccactatagaatggccctgtatgcacaagtctccgtccgaagagcttaccagat	1247
Db	1176	GTACACAGCCCAACGAGATGTGGCCCTGATGTGCACAAAGGCCCTGTGMAAGTGATACGAGCT	1235
QY	1248	cactctcccaagggcaccgccaacagccagtgatgtggcagbtgcacacttgcacccctgcg	1307

```

|||||
Db 1236 CATTCTCCACAGGCGCACCGCCACAGCAGTGATGGAGTGCACAACTCAACCTTCGCG 1295
Qy 1308 ggcctgaagaatgtaacccgccccagagcccaagggccacaccgcccgaagaagcgaa 1367
Db 1296 GCGCGAAGACATGTACATGGTCCAGAGCCACAGTGGCCACCAAAAAGACGGGCA 1355
Qy 1368 gaactccaggtctttagaaccctcagtggtgactgaagcagc---ggtggcagag 1423
Db 1356 GATCTCAGAGTCTTTAGAAATCCCTACGTGTGGACTAAGGCGACAGCTGTGGCCAGA 1415
Qy 1424 aagagcggtcaggtatggagagggccctgagagacctgccccgggcaaggaactctcag 1483
Db 1416 GGAAGCAGGTTGATTTAGAGAGGGCCCTTGAGACCTGTTGGTGAAGGACACACAGGG 1475
Qy 1484 gctcct-----cctccctgacagggcccaagaacatggtgccagatgtg 1530
Db 1476 AGCCCATGCTATAGGCTCCCTTCCTTCTTACTGTTCAAGGTAAGCTTATGCTCAAGG 1535
Qy 1531 aagggcctcctcctctccagtggttggtggtcagtggtgtcccccacccctcctc 1590
Db 1536 TCAGAGCCCTTGATCTGCTGATGTTGTTGGTGTCCATGAGGAGCCCTCCCA---CC 1592
Qy 1591 agtgttggtgagtcagagagcaaccacccagccctcctgcaagatacactcgcgtcag 1650
Db 1593 AGTATTTGTGAGTCCAGAGGCGCCCTGCTGTTTCATGCGCAGGCGCACCTGTGATCTA 1652
Qy 1651 actccagcaaatagtgctc---gggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1707
Db 1653 ACTTATTCACAAATAGCTTGTGCTTGGGATGTGGCTGTGTTCTATGCTTACGCTCCGG 1712
Qy 1708 agatctctgcaa-cctaaagagacttccacagggcctcag---gacctgacttctcctc 1763
Db 1713 TGACACCTATTAAGCCTCATTTGAGCTCAGAGAGATGCGCTTAAAGCTGATCTGCTCTC 1772
Qy 1764 tgtgaggaacaagaggtgctcaataataacatcttctcttatt 1806
Db 1773 TGTGAGGAGTAAAGGCTGCTTAAATATCATTTCTGCTTTATT 1815

RESULT 2
AL545009 903 bp mRNA linear EST 16-FEB-2001
LOCUS AL545009 LTI_NFL006.PI2 Homo sapiens cDNA clone CS0D1013YL0 5
DEFINITION prime, mRNA sequence.
ACCESSION AL545009
VERSION AL545009.1 GI:12877490
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 903)
AUTHORS Li, W. B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1. 903
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1013YL0"
/clone_1id="LTI_NFL006.PI2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(OT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,"

```

```

a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 132 a 302 c 263 g 206 t
ORIGIN

```

```

Query Match 48.7% Score 885.6; DB 9; Length 903;
Best Local Similarity 99.6% Pred. No. 2.2e-177;
Matches 888; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

Qy 49 ggaacccaacagagcgctgctgagagcaagatggcattccacaagacctgtgtatg 108
Db 11 GGGACCCACACAGACCTCGGCTGGAGCCAGAGATGGCATCCAAAGCCTTGGTATG 70
Qy 109 tgcctgagactgctctctctctctctctctcccaaggggcttggcccaagtcacacc 168
Db 71 TGCTGGGAGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 130
Qy 169 ggcctgagcaagagccttaacccctgtactaacaactgtgtgacgctctggggcgtg 228
Db 131 GCTTGACACCAAGGCTTCAACCCCTGTACTACAACTGTGTGACCGCTTGGGGCGTGG 190
Qy 229 ggcattgctctggaagcgctgctgtgagcgagatgtcacacgcttgtctcaacatc 288
Db 191 GGATCTGCTTGAGAGCGCTGCTGAGCGGCGGCGATTTGCACACCTTTGTGCTCACCATC 250
Qy 289 atcctgtgtcagacgctcctcctctctctctctctctctctctctctctctctctct 348
Db 251 ATCTGCTGAGCGACGCTCCCTCTTGTGAGAGACACCAAGAAAGGAGCTGTGGGAGAC 310
Qy 349 caggtatctctctctctctctctctctctctctctctctctctctctctctctctct 408
Db 311 CAGGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 370
Qy 409 aagcccgactctccacgtctgctctctctctctctctctctctctctctctctctctct 468
Db 371 AAGCCCGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 430
Qy 469 tgcctctctctctctctctctctctctctctctctctctctctctctctctctctctct 528
Db 431 TGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 490
Qy 529 gggcccgaggctgggtgactcttcaactgtgctctgtctgtgacccgtgtgaggtcatc 588
Db 491 GGGCCCGGGGCTGGGTGATCTTCACTGTGCTGTGCTGTGACCCCTGTGAGAGGTATC 550
Qy 589 atcaatagagatggtctatcaatcaacccctgttctggagcaatggagggcgccctcag 648
Db 551 ATCAATACAGAGTGGCTATCTATCTACCTGTTCCGGGCAAGTGGAGGCGGCCCTCAG 610
Qy 649 ggaacacagcagcagagctgtgacgttgcctccctctctctctctctctctctctctct 708
Db 611 GGAACACAGCAGCAGCGGTGGGCGTGGCGCTCCCTCTGTGCATGCCAACATGACATTT 670
Qy 709 gtcatgcaactcatctacgtactgtctgtctgtctgtctgtctgtctgtctgtctgtcc 768
Db 671 GTCAATGAGCAGCATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCT 730
Qy 769 ggcctgtgtgacgtctaaagcgctgagcaatgagcaatgagcaatgagcaatgagcaat 828
Db 731 GCGCTGTGGCGCTTAAGGCGCTGGGCTAATCATTTGGGCTTTGTGTCTCTACACACA 790
Qy 829 gccacctcgttgcacatatggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 888
Db 791 GCGACCTCGCTTGCCATATGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 850
Qy 889 cacaacagctccacactggatgatacccaacgctgtgacatgcctgtgcgcgcgcgcgc 940
Db 851 CACAACAGTCCCACTGGGATGATCCCAACGCTGCGCATTCGCGCGCAA 902

```

RESULT 3
AL570913/c 955 bp mRNA linear EST 16-FEB-2001
LOCUS AL570913 LTI_NFL006.PL2 Homo sapiens cDNA clone CS0D1013YL10 3
DEFINITION prime, mRNA sequence.
ACCESSION AL570913
VERSION AL570913.1 GI:12927686
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 955)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source
1..955
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1013YL10"
/clone_1b="LTI_NFL006.PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site 1: NotI. 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email: fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 175 a 275 c 301 g 200 t 4 others
ORIGIN

Query Match 48.5%; Score 882.8; DB 9; Length 955;
Best Local Similarity 98.7%; Pred. No. 8.8e-177;
Matches 918; Conservative 3; Mismatches 5; Indels 4; Gaps 3;
QY 806 gggcttctgctcctccaccacagccctccgttcgcatatggtgtgtgagatcgca 865
|||||
Db 936 GGGCTTTGTGTCT-CCTCACACAGCCACCTCGTTGCCATATGGGTGGTGGATCGTCA 878
QY 866 tgtatcttaccgacacacagacacacagtcaccacttgatgacccacgctgcca 925
|||||
Db 877 TGTATACTTACGGCAACAGCAGCAACAGTCCACCTGGGATGACCCGACGCTGGCA 818
QY 926 tgcgcctcgcgcgaatgctctggtgccttccttccttcacgtcat-ccccgaagctcc 984
|||||
Db 817 TCGGCTTCGCGCCCAATCCTGGGCTTCGTCTTCTACGTCAATCCCGAGGCTTCC 758
QY 985 caggctgaccagatccagccacagaaagctaccacggggagcatctaccaccgaggggc 1044
|||||
Db 757 CAGGTGACCAAGTCCAGGCCAGAGCAAAAGCTAACAGGGGAGACATGTACCCACCCGGGGC 698
QY 1045 gttgggctatgagaccatctctaaagacagaaggttcagagcatgtctgtgagaacaag 1104
Db 697 GTGGGCTTGTGAGACCATCTTGAAGAGCAGAAAGGTCTAGAGCATGTTCTGGAGAACAG 638
QY 1105 gcccttccatgagatgagcgggtgtgacagtaagagccggtgtgcaacatacagcggtac 1164
|||||
Db 637 GCGTTTTCATGAGATGACCCGTTTCACCTAAGAGCCGCGTCTCACCATACAGCGCGTAC 578
QY 1165 aatggagcagcgtcagcagcgtgtgtaccagccacatgagatgagccctgatacacaagt 1224
|||||
Db 577 AATGGGACAGCTGCTGACCAAGTGTGTACCAAGCCACTGAGATGGCCCTGATGACAAAGTT 518

QY 1225 ccgtccgaagagcttaccacatcctccaccagggccaccgccaacagccaggctgag 1284
|||||
Db 517 CCGTCGAAGAGGCTTACGACATCATCTCCACGGGGCCACCGCCACAGCAGGATGATG 458
QY 1285 ggcagtgccaatctgacctgtggtctgaagacatgtactgtgccccagagcccaagcg 1344
|||||
Db 457 GCGAGTGCACAACTCGA-CGTGCGGCTGAAAGACATGTACTCGGCCAGACCCAGAGG 400
QY 1345 gccacacggccgaaagacgcaagacactctcagctctttagaaccctcagctgtggag 1404
|||||
Db 399 GCCACACCGCCCAAAAGACGGCAAGAACTCTCAGGCTTTAGAAACCCCTACGTGGGAC 340
QY 1405 tgaagtcagcgtgtgcgagagagcgttcgagattgtggagagggccctagagaccctg 1464
Db 339 TGAGTCACGCGGTGGGAGAGAGAGCGGTCGATTTGGGAGAGGCCCTTGAGACCTGGCC 280
QY 1465 cgggcaagggactctccaggtccctccctccctccctccctccctccctccctccag 1524
|||||
Db 279 CGGGCAAGGGACTCTCAGGCTCTCTCCCTCCCTCCGACGCCACACATGTGCCAG 220
QY 1525 atgtgaagggcctccctcctctcgtccagtgattgtgggtgtgcatgtgtccaccac 1584
|||||
Db 219 ATGTGGAAGGCGCTCTCTCTCTGCGCAGTGTGGGTGGTGTGATGCGTGTCCACCCA 160
QY 1585 ctccctcagtgattgtgagtcgagagagcacaacccagcctctctgcagatcactcggc 1644
Db 159 CTCTCAGTGTGTGTGAGTGCAGAGARCCAAACCCACTCTTACACGAGAMCAGCTCGGC 100
QY 1645 ggtcacctccagccaatagtgtctctgggtgtgtgtgtgtgtgtgtgtgtgtgttc 1704
|||||
Db 99 GGTCACTCTCAACCAATAGTGTCTCGGGGTGTGTGTGTGTGTGTGTGTGTGTGTCTC 40
QY 1705 tggagattccttcaacctcaagagacttc 1734
|||||
Db 39 TGGAGATTCCTGCAACCTCAAGATTTGCC 10

RESULT 4
AL577178/c 906 bp mRNA linear EST 16-FEB-2001
LOCUS AL577178 LTI_NFL006.PL2 Homo sapiens cDNA clone CS0D1084YM03 3
DEFINITION prime, mRNA sequence.
ACCESSION AL577178
VERSION AL577178.1 GI:12940052
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 906)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source
1..906
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1084YM03"
/clone_1b="LTI_NFL006.PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site 1: NotI. 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email: fliang@lifetech.com URL :

FEATURES	source
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	
ORGANISM	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
RESULT	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	

[illegible]

```

DEFINITION AL546406.LTI.NFL006.PL2 Homo sapiens cDNA clone CS0D1031YJ06 5
prime, mRNA sequence.
ACCESSION AL546406
VERSION AL546406.1 GI:12879488
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 851)
AUTHORS Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source
1. 851
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CS0D1031YJ06"
/clone="LTI.NFL006.PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 127 a 281 c 245 g 196 t 2 others
ORIGIN
Query Match 46.0%; Score 837.6; DB 9; Length 851;
Best Local Similarity 99.5%; Pred. No. 3.2e-157;
Matches 848; Conservative 2; Mismatches 1; Indels 1; Gaps 1;
QY 176 gccaaagcccaaacccctgactacaaacctgtgtgacccgtctcggggcggtgggacatcg 235
Db 1 gccaaagccctacacccctgactacaaacctgtgtgacccgtctcggggcggtgggacatcg 60
QY 236 tccctgagagccgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 295
Db 61 tccctgagagccgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 120
QY 296 tggcagacccctcccttgtgtgaggaacaagaagagagcctgtgtgtgtgtgtgtgtgtgt 355
Db 121 tggcagacccctcccttgtgtgaggaacaagaagagagcctgtgtgtgtgtgtgtgtgtgt 180
QY 356 tcttcctcttggggagacccctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 415
Db 181 tcttcctcttggggagacccctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 240
QY 416 actctccacactgtgactcctcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 475
Db 241 actctccacactgtgactcctcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 300
QY 476 ctgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 535
Db 301 ctgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 360
QY 536 gggagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 595
Db 361 gggagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 420
QY 596 caagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 655
Db 421 caagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 480

```

```

QY 656 gccagcagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 715
Db 481 gccagcagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 540
QY 716 caactcatcactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 775
Db 541 caactcatcactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 600
QY 776 gtggcgctacaagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 835
Db 601 gtggcgctacaagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 659
QY 836 cccgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 895
Db 660 cccgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 719
QY 896 gtccacactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 955
Db 720 gtccacactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 779
QY 956 tccctcttctacgtatcccccgaagcttcccaagtgaccgaagtcagccagagcaagct 1015
Db 780 tccctcttctacgtatcccccgaagcttcccaagtgaccgaagtcagccagagcaagct 839
QY 1016 accagggggagca 1027
Db 840 accagggggagca 851
RESULT 8
AL571754/c 842 bp mRNA linear EST 16-FEB-2001
LOCUS AL571754.LTI.NFL006.PL2 Homo sapiens cDNA clone CS0D1031YJ06 3
DEFINITION AL571754.LTI.NFL006.PL2 Homo sapiens cDNA clone CS0D1031YJ06 3
prime, mRNA sequence.
ACCESSION AL571754
VERSION AL571754.1 GI:12929365
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 842)
AUTHORS Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source
1. 842
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1031YJ06"
/clone_lib="LTI.NFL006.PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 147 a 248 c 276 g 166 t 5 others
ORIGIN
Query Match 43.6%; Score 792.6; DB 9; Length 842;
Best Local Similarity 98.1%; Pred. No. 1.1e-157;

```

	Matches	828: Conservative	4: Mismatches	9: Indels	3: Gaps	3:
QY	871	acttcagcgaacagacagacacacagtcctccactcgtggatgacccacagctcgtccatcgc	930			
Db	842	ACTTCACGGGCAACACAGACACMACAGTCCACCTGGATGATCCCAACGCTGGCCATCGCC	783			
QY	931	ctcgccgcgaatagctctgtagccttgctcctctctagtcattcccgagttccccaagt	990			
Db	782	CTCGCTGGCCAAATGCGCT-GGCTTTCGTCTCTCTACGTATATCCCGA-GTCTCCAGGTG	725			
QY	991	accaagctccagcccaagagaagaactccagaggagacatgtaacccaccggggcttgagc	1050			
Db	724	ACCAAGTCCAGGCCACAGAGCAAAAGCTACCAAGGGGSACATGTACCCACCCGGGGCTGGCC	665			
QY	1051	tatgtagaccatctcgtaaagagcagaagaggtcagagcatgtctgtgagagaacaagccttt	1110			
Db	664	TATGAGACCATCTCTGAAGAGCAGAAAGGCTAGACATATGTTGTTGGAGAACAAAGCCTTT	605			
QY	1111	tccatcgtatgtagccggtctgtagcattaaggccggtgttaccatacagcgggttaacatgg	1170			
Db	604	TCCATGTGAATGAGACCCGCTGCCACGTAAAGGCCGGGTCTACCTTACAGGGGTACATGGG	545			
QY	1171	cagctgcctgaaccagctgtgttaaccagcccaacttgagatcggccctgatatcaaaattccgttc	1230			
Db	544	CAGCTGCTBACCAAGTGTGTACCAAGCCCACTGATGATGGCCTTATGCACAAATTCCTCGTCC	485			
QY	1231	gaaggagactttagaatactcatctctccacagggccacccgccaacagccaggtgaatggcagt	1290			
Db	484	GAAAGCGGCTTACAGACTCATCTCTCCACAGGGCCACCGCCACAGCCAGGATGTGGGCAAGT	425			
QY	1291	gccaaactcgaccctcgcgggctctgaagaatgtctctcggcccaagaccacaggcggccaca	1350			
Db	424	GCCCACTGCA-CCTGCGGGCTGAAAGACATGTACTGGGCCACAGCCMCAGCGGCGCACAA	366			
QY	1351	ccggcgaaagagcggcaagaactctcaggtcctttagaaaccttaacaccttgcgtgtgagactgagtc	1410			
Db	366	CCGCCGAAGAGACGCAGAGNACTCTCAGAGTCTTTTAAAGACCCCTACGTGTGGAGACTGAGTIC	306			
QY	1411	agcggatggcgaagagagagcggctcggatatttggagagggcccttgaaggacctcgtgcccgggca	1470			
Db	306	AGCGGTGGCGAAGAGAGGGCGGTGGGATTTTGGGAGAGGCCCTGAGGACCTGGCCCCGGGCA	246			
QY	1471	agggaactctcagactctcctctcctccctcgtgcaggcccaagaacaatgtgccagaatgtgg	1530			
Db	246	AGGGACTCTCCAGGGCTCCTCTCCCTCCCTGGCCAGGCCAAGCATGTGGCCACAGATGTGG	186			
QY	1531	aaggggctcccccctctctcgtcaggtgttggtaggtgtcaagggtgtcccaaccaactctctc	1590			
Db	186	AAGGGCTCTCCCTCTCTGCCAGTGTTTGGTGGGTGTCAATGGGTGTCCCAACCACTCTCTC	126			
QY	1591	agtgcttctgtgagtcgagagagcacaaccccaagcctctctgcagaatcactctggcgtaac	1650			
Db	126	ACTGTATTGTGGAGTGTGAGAGGACCAACCCAGCCTCTCTGCAGAGTACCTCGGGGTAC	66			
QY	1651	actccagcccaaatagtgctctcgtgggtgtgtgctcgtggcagcgctatgtttctcttgaga	1710			
Db	66	ACTCCAGCCAAATAGTGTCTCTGGGGTGTGTGCTGGCGCCCTATGATTTCTCTGGAGA	6			
QY	1711	tccc 1714				
Db	5	TCCTC 2				
RESULT	9					
LOCUS	BG829477					
DEFINITION	60276377461 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:489018 5',	956 bp	mRNA	linear	EST 22-MAY-2001	
ACCESSION	BG829477					
VERSION	BG829477.1	GI:14177064				
KEYWORDS	EST.					
SOURCE	human.					
ORGANISM	Homo sapiens					

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	(bases 1 to 956)	NIH-MGC http://mgc.nci.nih.gov/ .	National Institutes of Health, Mammalian Gene Collection (MGC)	Unpublished (1999)
	Contact: Robert Strausberg, Ph.D.	Email: cgabds-remail.nih.gov		
	Tissue Procurement: ATCC	CDNA Library Preparation: Ling Hong/Rubin Laboratory	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)	DNA Sequencing by: Incyte Genomics, Inc.
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at:	http://image.llnl.gov	Plate: L1CM1790	row: c column: 19
	High quality sequence stop: 770.	Location/Qualifiers:		
FEATURES	Source	1..956		
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/clone="IMAGE:4899018"			
	/clone_lib="NIH-MGC.42"			
	/tissue_type="epithelioid carcinoma cell line"			
	/lab_host="DH10B (phage-resistant)"			
	/note="Organ: pancreas; Vector: pOTB7; Site-1: XhoI; Site-2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC Library. 1"			
BASE COUNT	147 a	300 c	293 g	216 t
ORIGIN				
Query Match	42.6%;	Score 775.2;	DB 10;	Length 956;
Best Local Similarity	96.1%;	Pred. No. 5.5e-154;		
Matches 848;	Conservative 0;	Mismatches 28;	Indels 6;	Gaps 5;
QY	17	cagccggaagtaagagtcgctcagcctctgagggagcccaacagagcttgctggag	76	
DB	2	CAGCCGGGAAGTACGATGCTGCTCAGCCCTGGAGGACCCACAGAGCTGCTCTGTTCC	61	
QY	77	caagaatgcatccacaagaagcttgatgatctgctggagctgctctctctctgtcc	136	
DB	62	CCAGGATGGCCATCCACCAAGCCTTGATGATGCTTGAGACTGCTCTCTCTGTTCC	121	
QY	137	caggggctctgagcccaaggacatgtccaccacccgctcagagccaaaggctcaacccctgt	196	
DB	122	CAGGGGCTCTGGGCCCCAGGGCCATGTCACCCGGCTGCAGCCAAAGGCTCAACCCCTGT	181	
QY	197	actacaacctgtgtgacgcgtcttgaggcgctggggacgtctctggaagcgctgtgagg	256	
DB	182	ACTACAACCTGTGTGACCGCTCTGGGGGCTGGGGCATGCTCTGGAGGGCCGTGCTGGGG	241	
QY	257	cgggatgttcaacacgcttgtgtcaccacatcatctgttgccagagccctccctctgtgc	316	
DB	242	CGGGCATTTGTCACCAAGCTTGTGTGCTACCATATCCTGGTGCCACAGCCCTCCCTTTGTGC	301	
QY	317	aggacaaccaagaacgagcctctctgaggagcccaagatctctctctctctgaggaccctg	376	
DB	302	AGGACACCAAGAAAGGACCTGCTGGGGACCCAGATATTCTTCTTCTGGGGACCTGG	361	
QY	377	gacctctctgcctgtgttctgctcgtgtgtgtaagcccgactctcacaactgtgctctc	436	
DB	362	GCCTCTTGCTGCTGTTGGCTGTTGGTGTGAAGCCGACATTCACACTGTGCTCTGC	421	
QY	437	ggcgctctctcttgggggtctgtcgcacatctgctctctgtctgtcggggtacgctc	496	
DB	422	GGCGCTCTCTTTGGGGTTCGTTTCGCACATGCTTCTCTTGTCTGGGGGCTCACGCTC	481	

QY 497 tggccctcaactctctgcccgaagaacacagggcccggggtgtgattcttcaatg 556
 Db 482 TTGGCCCACTTCTCTGCCCCGGAAGAACACGGGCCCCGGGCTGGTATCTTCACTG 541
 QY 557 tggctctgtctgtacccctgttagagtcatacaatacagagtggtgtatcaacc 616
 Db 542 TGGCTCTGCTGCTGACCCCTGGTAGAGTTCATCATCATATACAGAGTGGCTGATCATCACCC 601
 QY 617 tggttcgggagcagtgtggagagggccctcaagggaacagcagcagctgtggcgttg 676
 Db 602 TGGTTCGGGGCAGTGGGAGGCGG-CCTCAGGGCAACAGCAGGCGGCGGCGCTGG 660
 QY 677 cctcccccctgtccatgcacaacatgactgttcatgagccatcatcagctacgc-tg 735
 Db 661 CTTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
 QY 736 ctgctgtgtgtgtctctctgtggggcctgtgcccgtgtgtgtgtgtgtgtgtgtgt 795
 Db 721 CTGCTGCTGGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
 QY 796 cgttagagatgt 853
 Db 781 GTAAGGCAATGGGCTGTGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
 QY 854 tgt 893
 Db 841 TGT 882

RESULT 10
 AL578937/c 846 bp mRNA linear EST 16-FEB-2001
 LOCUS AL578937 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DK005YN05 3
 DEFINITION AL578937 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DK005YN05 3
 prime, mRNA sequence.
 ACCESSION AL578937
 VERSION AL578937.1 GI:12943490
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (Bases 1 to 846)
 AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: sequef@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
 Source
 1. 846
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0DK005YN05"
 /clone_lib="LTI_NFL006_PL2"
 /tissue_type="placenta"
 /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifestech.com URL : http://fulllength.invitrogen.com"
 BASE COUNT 148 a 250 c 269 g 175 t 4 others
 ORIGIN

Query Match 41.8%; Score 761.2; DB 9; Length 846;
 Best Local Similarity 97.8%; Pred. No. 4.9e-151;

Matches 831; Conservative 4; Mismatches 8; Indels 7; Gaps 6;
 QY 874 taaggcaacaagcagcaacaacagtcacacctggatgaccccaagctggccatgcctc 933
 Db 846 TAGSCAACAAAGCAGCAACAAAGTCCACCT-GGATGACCCACAGCTGGCCATGCCCCCTC 788
 QY 934 gccgcacaatgcttgccctctctcctctctcagtcacatccccc--gaagctcccaagtga 991
 Db 787 GCCCCCAATGCTTGGG-CYTGCTCTCTTTTACGTATCCCCCAGAGCTCTCCCAAGGTGA 729
 QY 992 ccaatgccag-cccagagcaagctacccaaggaggacatgtaaccccaacccgggagtgagc 1050
 Db 728 CCAAGTCCAGCCCGAGCAAGCAAGCTACAGGGGAGCATGTACCCACCC-GGGGCTGGGC 670
 QY 1051 tatgagacatccctggaagagcaagagtgtaagagcatgttctgtgtgagaaagagcttt 1110
 Db 669 TATGAGACATCTCTGAAGAGCAGAGAGCTGAGAGCATGTTCGTGGAACAGCCCTTT 610
 QY 1111 tccatgatatgagcagtgtagcagcttaagagcagctgtacacatacagcggtacaaatgag 1170
 Db 609 TCCATGATGAGCCGGTTCGACTAGAGGCCGGTGTACCATATCAGCGGCTACAAATGG 550
 QY 1171 cagctgtacacagtggtgtacacagccactgagatgagcctgtatgcaaaagtctcgtcc 1230
 Db 549 CAGCTGTGACCAAGTGTGTACAGCCCTGAGATGGCCCTGATGCAACAAAGTTCCGTCC 490
 QY 1231 gaagagcttacagcatcatctctcccaaggagcaccgccaacagcaggtgtatggcagt 1290
 Db 489 GAAAGAGCTTACGACATCATCTCTCCAGGGGCCACCCGCAACAGCAGCATGATGGCGGT 430
 QY 1291 gcccaactgcaccctgcgagctggaagacatgtaactcgtgcagcagcagcagcgagcaca 1350
 Db 429 GCCAACTGCA-CTGCGGGCTGGAAGACATGTACTGCGCCAGAGCCACAGCGGCCACCA 371
 QY 1351 ccgcgcgaagaagcgaagaactctcaagctcttagaaacccctacgtgtgtgagctgagtc 1410
 Db 370 CCGCCGAAGAGCGGCAAGAACTCTCAGGTCTTTAGAAACCCCTAGTGTGGAGCTGAGTCT 311
 QY 1411 agcggtgagcagagagagcagctgagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1470
 Db 310 AGCGGTGCGGAGGAGAGTGGCGGATTTGGGAGGGCCCTGAGAGACTGGCCCGGGCA 251
 QY 1471 agggactctcagagctctctctccctctgcagcagcagcagcaaatgtgtcccaatgtgt 1530
 Db 250 AGGCACTCTCAGAGCTCTCTCTCCCTCTGCGAGGCCAGCAACATGTGCCCCAGATGTGG 191
 QY 1531 aaggacctcctctctgtgcagtgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1590
 Db 190 AAGGGCCCTCCCTCTGCTGCAAGTGTGTGGGTGGGTGTGATGTGGTGTGCCACCACTCCTC 131
 QY 1591 agt 1650
 Db 130 AGGTGTTGTGAGTGTGAGAGAGCCAAACCCACCTCTCCAGGATCAACCTCGGGGTGAC 71
 QY 1651 actcagcacaatagtgtctcgcgggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1710
 Db 70 ACTCCAGCAATATGTGTTCTCGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 11
 QY 1711 ttctgtcaac 1720
 Db 10 TTCTGTCAAC 1

RESULT 11
 LOCUS B1769602 873 bp mRNA linear EST 25-SEP-2001
 DEFINITION 603054983F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5204373 5',
 mRNA sequence.
 ACCESSION B1769602
 VERSION B1769602.1 GI:15761180
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 873)
 AUTHORS NIH-MGC
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LRAM1512 row: f column: 22
 High quality sequence start: 3
 High quality sequence stop: 821.
 Location/Qualifiers

FEATURES

source

1. .873
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5204373"
 /clone_lib="NIH_MGC_122"
 /lab_host="DH10B"
 /note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 24 week female lung, 16 week female spleen, and 20-22 week male spleens. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH-MGC Library."
 BASE COUNT 126 a 285 c 261 g 201 t
 ORIGIN

Query Match 41.1%; Score 747.8; DB 10; Length 873;
 Best Local Similarity 95.2%; Pred. No. 3,4e-148;
 Matches 837; Conservative 0; Mismatches 32; Indels 10; Gaps 6;

Qy 40 cagcctgtaggagaccacacagagcctgctgtaggagccagatgcatcacacaagcc 99
 |||||||
 Db 1 CAGCCTGAGGAGGACCCACACAGAGCTGG-CTGGGAGCCAGATGGCCATCCACAACCC 59
 Qy 100 ttgtgtatgtgctgtaggagctctctctctgttccaggggctggtgcccaggccat 159
 |||||||
 Db 60 TTGTGTATGTCTCTGTGGGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 117
 Qy 160 gtccacacccgctgtaggagcagagcctcaacccctgtactaaccctgtgagagcct 219
 |||||||
 Db 118 GTCCACACCCGCTGAGCCAGCCCTCAACCCCTGTACTACACCTGTGTAGCCGCTCT 177
 Qy 220 ggggcgtgtaggagcctgctgtaggagcgtggtggtggtggtggtggtggtggtg 279
 |||||||
 Db 178 GGGGCGTGGGCGATGCTCTGTGAGGCGCTGGCTGGGGCGGCGATGTACACAGTTTGTG 237
 Qy 280 ctacacatcatcctgtaggagcagcctccctctgtgtaggagacacaaagaagagcctg 339
 |||||||
 Db 238 CTCACACATCATCTGTGTGGCAGCCTCCCTTTGTGACGACACCAAGAAAGAGAGCTG 297
 Qy 340 ctggggagaccggtatctctctctctgtaggagccctgagcctctctgctgctgttggc 399
 |||||||
 Db 298 CTGGGGAGCCAGGATTTCTCTTGTGGGAGCCCGGGGCTCTTCTGTGCTGTGTGCTG 357
 Qy 400 tgtgtgtaggagccagactctccacccgtgacctctgagcgtctctctcttggagttctg 459
 |||||||
 Db 358 TGTGTGGTGAAGCCCGAGATTCTCACACTGTGCTGTGGCGCTTCTTGTGGGTTTCTG 417
 Qy 460 ttgcacatctctctctctgtctgtaggagcagcctcttggcctcaactctctggcccg 519
 |||||||

Db 418 TTGGCATCTGCTCTCTCTGTGTGTGGGCGCTACAGCTTTTGCCCTCACTTCCTGGCCGG 477
 Qy 520 aagacacagggcccccggggctgggtgtagtcttcaactgtgctgtctgtacacctgtgta 579
 |||||||
 Db 478 AAGAACACAGGCCCCGGGGCTGGGTGATCTTCACTGTGGCTCTGCTGACCTGGTA 537
 Qy 580 gaggatcatcatatagagtagtctgtatcatcaccctggtctggggagctggcagagc 639
 |||||||
 Db 538 GAGGTCAATCAATATACAGAGTGGCTGATCATACACCTGTGTTGGGGCAGTGGCAGGCG 597
 Qy 640 ggcctcagggcagaagcagcagcagcgtggtggtggtggtggtggtggtggtggtg 699
 |||||||
 Db 598 GG-CCTCAGGGCAGACAGGAGCGAGCGAGCGCGCGCGCTCCCTCTGTGCGCATGCCAAC 656
 Qy 700 atgagc-cttgcattgagacatcatcagtcagtcagtcagtcagtcagtcagtcagtcag 758
 |||||||
 Db 657 ATGACCTTTTGTGATGACATCATATACGTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 716
 Qy 759 ggcctgagccgcccctgtgtgtagcgtacaaagcgtgtagcgtgtagcgtgtagcgtgtagc 818
 |||||||
 Db 717 GGCCTGGGCGGCGCTGTGTGGCGGCTACCAAGCGCTGGGCTAAGCATGGGGTCTTTGTGCT 776
 Qy 819 cctcaacacagccacccctcgtgtgcatatggtgtg--tgtagtcgtcatgtatacttac 876
 |||||||
 Db 777 CCTCACACACAGCCACCTCCGTTGGCATATGGGTGGGTGCTGTGATGCTGTATCTTAC 836
 Qy 877 ggcacaaagcagcagaacagtcaccctggatgagcccg 915
 |||||||
 Db 837 GGCAAAAGCACCAACA---GTCCACTGGGATGACCCC 872

RESULT 12
 BI834606 820 bp mRNA linear EST 04-OCT-2001
 BI834606
 LOCUS 603089872P1 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5228800 5',
 DEFINITION mRNA sequence.
 ACCESSION BI834606
 VERSION BI834606.1 GI:15946156
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 820)
 AUTHORS NIH-MGC
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LRAM1575 row: p column: 17
 High quality sequence stop: 804.
 Location/Qualifiers

FEATURES

source

1. .820
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5228800"
 /clone_lib="NIH_MGC_120"
 /lab_host="DH10B"
 /note="Organ: pooled pancreas and spleen; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of spleen and pancreas from 28 yo male. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-2.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics

BASE COUNT 118 a 264 c 250 g 188 t
ORIGIN

Query Match 40.9%; Score 744.4; DB 10; Length 820;
Best Local Similarity 97.3%; Pred. No. 1.7e-147;
Matches 800; Conservative 0; Mismatches 16; Indels 6; Gaps 4;

QY 38 ctacagcttgagagagaccacagagccttgccctggagagccagatggtccatccacaag 97
DB 1 CTACAGCTTGAGAGGA-CCAAACCAAGAGCTGGCCCTGGAGCCAGATGAGCCATCCACAAG 59
QY 98 ccttgatgagtgctggagactgctctctctctctctctctctctctctctctctctctct 157
DB 60 CCTTGATGATGCTGCTGGAGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 119
QY 158 atgtccaccgagctgacgacgaagcctcaacccctgtactacaacctgtgtacgct 217
DB 120 ATGTCCACCCGCTGAGAGCCAGGCTCAACCCCTGTACTACACCTGTGTACCGCT 179
QY 218 ctggagcttgagagatcgtcttgagagcgtgtgctggagcgagcatgtcaacagtttg 277
DB 180 CTGGGGGCTGGGCGATGCTCTGGAGGCGGTGGGGCGGCGCATGTCTACACAGTTTG 239
QY 278 tgcacacatcatcctgtgtgagcagcctccctctgtgcaagagacagaagaagcgagcc 337
DB 240 TGCTCACCATCATCTGTGTGGCCAGCCTCCCTTTGTGACAGACACAGAAAGGAGCC 299
QY 338 tgcctggagaccagatct 397
DB 300 TGCTGGGAGCCAGGATATCT 359
QY 398 cctgtgtgtgagagccgagactctcaacgtgtcctctctctctctctctctctctctct 457
DB 360 CCTGTGTGTGAAACCCGAGCTTCTCCACCTGTGCTCTCTCTCTCTCTCTCTCTCTCTCT 419
QY 458 tgttcgcacatcgtct 517
DB 420 TGTTGGCCATCTGCTTCTCTGTGTGGGGGCTCAAGTCTTGGCCCTCAACTCTCTGGGCC 479
QY 518 ggaagaacacagcgcccgagctgggtgtatctcaactgtgtcctctctctctctctctct 577
DB 480 GGAAGAACAAGCGGCCCGGGGCTGGGATCTTCACTGTGTGCTCTGTGTGACCTGG 539
QY 578 taagagtcacatcaataacagatgctgtatcatcaacctgtgtctcggagcgagtgag 637
DB 540 TAGAGGTATCATCAATATACAGATGAGCTGATCATACCTGTGCTGGGGCAGTGAGG 599
QY 638 ggggacctcaaggggaaagagcgagctgtggcgctgtgcccctgtgtccatctgca 697
DB 600 GCGGACCTCAAGGCAAGGAGCGAGGCTGGGCTGGCCCTCCCTGTGCCATCGCA 659
QY 698 acatgagactgtcatgtgacatcatctacgtcactgtctgtctgtgtgtctctctctct 757
DB 660 ACATGGAATTGTGATGGCAATCATACGTACATGCTGTCTGTGGGTGCTCTTCG 719
QY 758 gggccttgagccgctgtgtgtgagcgctacaagcgctggcgtaagcatggagttctgtgc 817
DB 720 GGGCTGTGGCCCG-CCGTGTGTGGAGC-TACAAAGCGCTGGCGTAAGCTTAAAGTGC 777
QY 818 ttcctcacacagcagcctcgttgccatagtgtgtgtgtga 859
DB 778 TCCT--CACACAGGACTCCGTGCATATGATGTGTGTGA 816

RESULT 13
BI772730 933 bp mRNA linear EST 25-SEP-2001
LOCUS 603032860F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:520285 5',
DEFINITION mRNA sequence.
ACCESSION BI772730
VERSION BI772730.1 GI:15764308

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 933)
AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LHAM1508 row: 9 column: 16
High quality sequence stop: 782.

FEATURES
source
location/Qualifiers
1..933
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5202855"
/clone_1lb="NIH_MGC_122"
/lab_host="DH10B"
/note="Organ: pooled lung and spleen; Vector: pCMV-Sport6;
Site_1: NotI; Site_2: EcoRV (destroyed); RNA source
anonymous pool of 24 week female lung, 16 week female
spleen, and 20-22 week male spleens. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.4 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 026. Note:
this is a NIH_MGC Library."

BASE COUNT 141 a 303 c 276 g 213 t
ORIGIN

Query Match 40.2%; Score 732; DB 10; Length 933;
Best Local Similarity 92.9%; Pred. No. 7.7e-145;
Matches 866; Conservative 0; Mismatches 55; Indels 11; Gaps 9;

QY 55 caacagagcctggccttgagagcagatgagcatccacaagccttgatgtgctgt 114
DB 3 CAACGAGAGCTGG-CTGGAGACCAAGATGAGCATCCAAAGCCTTGATGTGCTGTG 61
QY 115 gg-actgct 173
DB 62 GGAAGCTGCT 119
QY 174 cagccaaagctcaacccctgtactacaacctgtgtacacgctctgggagctggagat 233
DB 120 CAGCCAAAGCCTCAACCCCTGTACTACACTGTGTACCGCTGTGGGGGTGGGAGAT 179
QY 234 cgtccctgagcgctgtgcttgagcgagcatgtcaacacgttctgtctacatcatcct 293
DB 180 CGTCTGAGGAGCGGTGCTGGGGGCGGCGATGTGCACACAGCTTGTGTCACACATCATCT 239
QY 294 ggtggcagcctccct 353
DB 240 GGTGGCCAGCTCCCTCTTGTGTGCAAGGACCAAGAAAGGAGCCGTGGGGAGCCCAAGT 299
QY 354 attctctctcttgagaccctggagcctctctctctctctctctctctctctctctctctct 413
DB 300 ATTCTCTCTCTGGGAGCCCTGGGCTCTTCTGCTCTGTTTGTGCTGTGTGTGAAAGCC 359
QY 414 cgaactctcaacctgtgtcct 473
DB 360 CGACTTTCACACGTGCTCTCGGCGCTCTCTCTTGTGGGTTCTGTGTGCGCATCTGCTT 419

OY	474	cttttgcttgaggaggtacagctcttggccctaaactccgcggcggaagaaccacgggac	533
Db	420	ctctttgtctggggggctcacgctttggccctcaacttctcgggccggamaaacacggggcc	479
OY	534	ccgggagctgagtgatccttcaactctgagctctgctctgacccctggtatagatcatcaaa	593
Db	480	ccggggctgggggagatcttccatcgtggcctctgctgacccctggtatagaggtcatcatcaa	539
OY	584	tacaaatggcctgatatcaacccttggtcttgaggaggtggagggcgggccctcaagggcaa	653
Db	540	tacaaatggcctgatatcaacccttggtcttgaggaggtggagggcgggccctcaagggcaa	598
OY	654	caggacgcgaagctgagccgttgagctctcccccctgctgacatcgcaacaatgagatctgtcat	713
Db	599	caggacgcgaagcct--ggccgtggcctccccctgtgacatcgcccaaatgagatctgtcat	657
OY	714	ggcactcaatctacgtcatcagctgctgctgctgagtgctcctcct--ggggagcctggccggccc	772
Db	658	ggcactcatcatcagctcatcctctcctctgctgggctccttctggggggccttggggccggcc	717
OY	773	tgtgtggccgctacacaagcagctgagctgaaacat--gggctctctgtgctcctccacacaagcc	831
Db	718	tgtgtggcccgctacacaagcagcctggcgttaagcattgggggcttgtgtgctcctcctcacacaaga	777
OY	832	acctcagcttgccatat--gggtgtgtgtgatactgcatatatacttaacttaagcaaacagcagc	889
Db	778	ccttcgcttggcatatttgggtgggggtggatggcattgattacttaacccggamaacagaca	837
OY	880	acaacagctcccaacccggg--atgacccaagcctgtgcatatcgccctcgccgcgaatgctgtt	948
Db	838	gaatacagttccacacgtcggaattgacccacacgcctggggccatcggcctcccccatttgcttggg	897
OY	949	ggcttcgctcctctctcaactgcatccccgaagt 980	
Db	898	ggcttctgctcccatataagatttaccggaggt 929	

RESULT	14
LOCUS	BG750651
DEFINITION	BG750651 776 bp mRNA linear EST 15-MAY-2001 60270851LPI NIH_MGC_43 Homo sapiens cDNA clone IMAGE:484930 5' , mRNA sequence.
ACCESSION	BG750651
VERSION	BG750651.1 GI:14061304
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo. 1 (bases 1 to 776)
AUTHORS	NIH-MGC http://mgc.ncl.nih.gov/.
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
JOURNAL	Contact: Robert Strausberg, Ph.D. Email: cgapbs@mail.nih.gov
COMMENT	Tissue Procurement: ATCC cDNA Library Preparation: Ling Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLCMI683 row: f column: 03 High quality sequence stop: 750.
FEATURES	Location/Qualifiers 1..776
source	

```

"organism":"Homo sapiens"
"db_xref":"taxon:9606"
"clone_image":"4844930"
"clone_lib":"NH.MGC.43"
"tissue_type":"normal pigmented retinal epithelium"
"lab_host":"DH10b (phage-resistant)"
"note":"Organ: eye; Vector: pORB7; Site_1: XhoI; Site_2:"

```

ECOR1: cDNA made by oligo-dT priming, directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAC(G) Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using 24c cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC library. |

Query Match	40.0%	Score 728.4	DB 10	Length 776
Best Local Similarity	99.0%	Pred. No. 4.2e-144		
Matches 754; Conservative	0	Mismatches 6	Indels 2	Gaps 2

QY	861	cgtaactgataacttaacttcggaacaagaacagacaaacagtcctaccactctggatagaccccaagct	920
Db	2	ccctcatgtatattcttccggaacaaacacagacacacagtcctaccactctggatagaccccaagct	61
OY	921	ggcaactgcacctcgcgcgcgaacatgcttg9gaccttgctctctctctactagtcattcccgaggt	980
Db	62	ggcccatgcgacctcgcgcgcgaacatgcttg9gaccttgctctctctctactagtcattcccgaggt	121
OY	981	ctcccaagtgacaaagtcaccagcccaagcaaaagctatacaaggggagaaatgtaaccacccg	1040
Db	122	ctcccaagtgacaaagtcaccagcccaagcaaaagctatacaaggggagaaatgtaaccacccg	181
OY	1041	ggcgctgggctctctgagacacatctctgaagaagacagaagggctcaagaatgcttgctggagaa	1100
Db	182	ggcgctgggctctctgagacacatctctgaagaagacagaagggctcaagaatgcttgctggagaa	241
OY	1101	caagcgctctccatgagatgaagcggcttgcaagctaaagagcgcggtgtccactataagcgg	1160
Db	242	caagcgctctctccatgagatgaagcggcttgcaagctaaagagcgcggtgtccactataagcgg	301
OY	1161	gtacaactgggacagctgtctgacagatgtgttaacagcccaactgagatgtgcccgtatgcaaa	1220
Db	302	gtacaactgggacagctgtctgacagatgtgttaacagcccaactgagatgtgcccgtatgcaaa	361
OY	1221	agttctctctccagagagagcttaagacatactctccacgggacggccacggccaagcaggt	1280
Db	362	agttctctctccagagagagcttaagacatactctccacgggacggccacggccaagcaggt	421
OY	1281	gattggcgactgacaaactcgaccctcgcggtgtaagaacatgtaactcctggcccaagacacca	1340
Db	422	gattggcgactgacaaactcgaccctcgcggtgtaagaacatgtaactcctggcccaagacacca	481
OY	1341	ggcggcacacacgcgcgaagaagcgaagaactcgaagtccttagaaccctctacgttgg	1400
Db	482	ggcggcacacacgcgcgcgaagaagcgaagaactcgaagtccttagaaccctctacgttgg	541
OY	1401	ggactgagatcgaacggttggcgaagaagaaggggtcggaattttggggaaggcccttgaagacctg	1460
Db	542	ggactgagatcgaacggttggcgaagaagaaggggtcggaattttggggaaggcccttgaagacctg	600
OY	1461	ggcccgaggcaagagactctccaggtctctctctcccccctggcgaaggcccaacaatgtgcc	1520
Db	601	ggcccgaggcaagagactctccaggtctctctctcccccctggcgaaggcccaacaatgtgcc	660
OY	1521	ccagaatgtgaaaggcctctccctctctccagatgttttgggttgatgtaatggtgtcccca	1580
Db	661	ccagaatgtgaaaggcctctccctctctccagatgttttgggttgatgtaatggtgtgtgt-ccca	719
OY	1581	cccaactctcaagtgttttgggttgatgtaaggaagacaaacccacagc	1622
Db	720	cccaactctctcaagtgttttgggttgatgtaaggaagacaaacccacagc	761
RESULT 15			
LOCUS	B1822698	1052 bp	mrna linear EST_04-OCT-2001
DEFINITION	60303603391 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5176977 5',		
mrna sequence.			

ACCESSION BI822698
 VERSION BI822698.1 GI:15934248
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 1052)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: rgs@bbs-rtmail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: L1A11441 row: a column: 10
 High quality sequence start: 2
 High quality sequence stop: 736.
 FEATURES
 source
 1..1052
 location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5176977"
 /clone_1b="NIH_MGC_115"
 /lab_host="DH10B"
 /note="Organ: pooled brain, lung, testis; Vector:
 pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
 source anonymous pool of 6 male brains, age range 23-27; 1
 male lung, age 27; and 1 male testis, age 69. Library is
 oligo-dT primed and directionally cloned (EcoRV site is
 destroyed upon cloning). Average insert size 1.8 kb,
 insert size range 1-3 kb. Library is normalized and
 enriched for full-length clones and was constructed by C.
 Gruber (Invitrogen). Research Genetics tracking code
 021. Note: This is a NIH_MGC Library."
 BASE COUNT 204 a 338 c 306 g 204 t
 ORIGIN
 Query Match 39.0%; Score 710.2; DB 10; Length 1052;
 Best Local Similarity 93.8%; Pred. No. 3.4e-140;
 Matches 806; Conservative 0; Mismatches 43; Indels 10; Gaps 6;
 Oy 45 tggaggagacccaacagagcctgagccttgagagccagagatggccatccacaagaagccttgt 104
 Db 1 TGGAGGGACCCAAACCAAGAGCCTGGCCTGGAGCCAGATGGCCATCCACAAAGCCTTGCT 60
 Oy 105 gatgtgacctggagactgctcc 164
 Db 61 GATGTGCTGGGAGCTGCT 119
 Oy 165 accgggctgacgcaaggcctcaaccctctactactaacaacctgtgtgacgcgtctggagc 224
 Db 120 ACCGGGCTGACGCAAGGCTCAACCCCTCTACTACAACTGTGTGACCGCTCTGGGGC 179
 Oy 225 gtggaggacgtctcctggagagccgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 284
 Db 180 GTGGGGCATTCCTCTGGAGGCGGTGGCTGGGGCGGCAATGTGCACCAACGTTGTGTCTCAC 239
 Oy 285 catcatcctggctggcacaagcctccctctgtgtcagagacacaaagaacgagcctgtggg 344
 Db 240 CATCATCCTGGTGGGCAACCTCCCTTGTGTGAGAGACAAAGAAAGGAGGCTGTGGG 299
 Oy 345 gacccaggtatctgtgtgtgt 404
 Db 300 GACCCAGGTATTCCT 359
 Oy 405 ggtgaagccgacactctcactgtgctctctcgtgagcctctctctctctctctctctgtctgc 464

Db 360 GGTGAAGCCGACATTCCTCCACCTGTGCTGGGGCTTCCTCATAGGGTTCTGTTGCC 419
 Oy 465 catctgtctctctgtctgtcgtgctcaagctctctgtccctcaactctctgtgcccggagaa 524
 Db 420 CATCTGCTTCCTTCTGTGGGGGCTCAAGTCTTGCCTCAACTCCCTGGGCCGGAAGAA 479
 Oy 525 ccacgggcccggggctgtgtgtatcttactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 584
 Db 480 CCACGGGCCCGGGGCTGGGTGATCTTCTGTGCTGTGCTGTGACCTGTGTAGAGT 539
 Oy 585 catcatcaatacaagatgtgctgatatcccttggtcgggagctggcgaagcgagcc 644
 Db 540 CATCATCATATACAGATGCTGATCATACCCCTGTTGGGGGCAATGGGCGAGGGGGCC 599
 Oy 645 tcaaggcaaa-cagcagcgagcgtgtggcgtgtgctccctctgtgcaatgcgcaa-catg 702
 Db 600 TCAGGGCAACCAGACGCGCAGGCTGGGCGCTCCCTCCCTGTGCTCAACCATG 659
 Oy 703 gactttgcatggagactatcatctatctgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 762
 Db 660 GACTTTGTCATGGACATCATCTACTATGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 719
 Oy 763 tggccggcctgtgtgtgctgtacaaagcgtgcg-----taagcatgggtctctgtgtgtgt 817
 Db 720 TGGCCGGCGCTGTGTGTGGCGGTACAAAGCGCTGAGCGTAAAGCATTTGGGGTCTATGTC 779
 Oy 818 tcttaaccac-agccacctcgtgtgcatat-aggatgtgtgtgtgtgtgtgtgtgtgtgtgtgt 875
 Db 780 TCTCTAACCAAGCAACCTCCGTAGCATATAGCGTGGCGTGGATGCAATGTATATACACT 839
 Oy 876 cggcaacaagcagcacaac 894
 Db 840 ACAGGAACACAGCAGCAC 858

Search completed: September 22, 2002, 16:13:13
 Job time: 5936 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2002, 14:39:52 : Search time 75.36 Seconds
(without alignments)
5928.972 Million cell updates/sec

Title: US-09-895-686-7

Perfect score: 1819
Sequence: 1 cgcgtcgcagccctaccagc.....ctttatcactcttaaaaa 1819

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued Patents.NA:*
1: /cgn2_6/prodata/2/lna/5A.COMB.seq:*
2: /cgn2_6/prodata/2/lna/5B.COMB.seq:*
3: /cgn2_6/prodata/2/lna/6A.COMB.seq:*
4: /cgn2_6/prodata/2/lna/6B.COMB.seq:*
5: /cgn2_6/prodata/2/lna/PCRTUS.COMB.seq:*
6: /cgn2_6/prodata/2/lna/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	191.6	10.5	2484	4	US-09-276-531-46 Sequence 46, Appl
2	187.4	10.3	1212	3	US-09-188-930-249 Sequence 249, App
3	118.8	6.5	311	3	US-09-188-930-4 Sequence 4, Appl
4	52.4	2.9	7218	1	US-08-232-463-14 Sequence 14, Appl
5	46.2	2.5	1035	3	US-08-733-837B-1 Sequence 1, Appl
6	45.4	2.5	1931	2	US-09-130-114-2 Sequence 2, Appl
7	45.2	2.5	6048	4	US-09-634-920-3 Sequence 3, Appl
8	43.6	2.4	4411529	4	US-09-103-840A-1 Sequence 1, Appl
9	41.6	2.3	1597	2	US-08-724-974A-1 Sequence 1, Appl
10	40.8	2.2	3384	2	US-08-687-289A-1 Sequence 1, Appl
11	40.8	2.2	3809	1	US-08-485-588-3 Sequence 3, Appl
12	40.8	2.2	3809	1	US-08-484-565-3 Sequence 3, Appl
13	40.8	2.2	3809	2	US-08-480-751-3 Sequence 3, Appl
14	40.8	2.2	3809	3	US-08-943-986-3 Sequence 3, Appl
15	40.8	2.2	3809	3	US-08-353-784-3 Sequence 3, Appl
16	40.8	2.2	3809	3	US-08-484-719B-3 Sequence 3, Appl
17	40.8	2.2	3809	4	US-08-546-998-2 Sequence 2, Appl
18	40.8	2.2	3809	4	US-08-484-159-3 Sequence 3, Appl
19	40.8	2.2	4465	1	US-08-180-195-1 Sequence 1, Appl
20	40.8	2.2	4465	1	US-08-477-329-1 Sequence 1, Appl
21	40.8	2.2	4465	2	US-08-475-458-1 Sequence 1, Appl
22	40.8	2.2	4465	3	US-08-980-400-1 Sequence 1, Appl
23	40.8	2.2	4465	4	US-09-583-459A-1 Sequence 1, Appl
24	40.8	2.2	4465	4	US-09-583-210-1 Sequence 1, Appl
25	40.8	2.2	4465	4	US-09-583-449A-1 Sequence 1, Appl
26	40.8	2.2	5006	1	US-08-485-588-2 Sequence 2, Appl
27	40.8	2.2	5006	1	US-08-484-565-2 Sequence 2, Appl

28	40.8	2.2	5006	2	US-08-480-751-2 Sequence 2, Appl
29	40.8	2.2	5006	2	US-08-943-986-2 Sequence 2, Appl
30	40.8	2.2	5006	3	US-08-353-784-2 Sequence 2, Appl
31	40.8	2.2	5006	3	US-08-484-719B-2 Sequence 2, Appl
32	40.8	2.2	5006	4	US-08-546-998-1 Sequence 1, Appl
33	40.8	2.2	5006	4	US-08-484-159-2 Sequence 2, Appl
34	40.8	2.2	5427	1	US-08-168-917-1 Sequence 1, Appl
35	40.8	2.2	5427	2	US-08-460-510-1 Sequence 1, Appl
36	40.8	2.2	5427	2	US-08-460-490-1 Sequence 1, Appl
37	40.8	2.2	5427	3	US-08-462-728-3 Sequence 3, Appl
38	40.8	2.2	5427	5	PCr-US92-00730-1 Sequence 1, Appl
39	40.8	2.2	5427	5	PCr-US92-00862-1 Sequence 1, Appl
40	40.6	2.2	289	4	US-09-007-005-17 Sequence 17, Appl
41	40.6	2.2	289	4	US-09-244-796-17 Sequence 17, Appl
42	40.6	2.2	3489	2	US-08-728-323A-1 Sequence 1, Appl
43	40.6	2.2	32207	2	US-08-770-379-20 Sequence 20, Appl
44	40.6	2.2	32207	4	US-08-757-669A-20 Sequence 20, Appl
45	40.6	2.2	32207	4	US-09-230-371A-20 Sequence 20, Appl

ALIGNMENTS

RESULT 1
US-09-276-531-46
Sequence 46, Application US/09276531
Patent No. 6183968
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
APPLICANT: Hillman, Jennifer L.
APPLICANT: Yue, Henry
APPLICANT: Reddy, Roopa
APPLICANT: Guegler, Karl J.
APPLICANT: Baughn, Mariah R.
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF GENES ENCODING
NUMBER OF SEQUENCES: 134
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/276, 531
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/079, 677
FILING DATE: March 27, 1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Lynn E. Murry, Ph.D.
REGISTRATION NUMBER: 42,918
REFERENCE/DOCKET NUMBER: PA-0008 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 2484 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: SYNORAT05

[illegible]

```

; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; ANTI-SENSE: NO
US-08-724-974A-1

```

Query Match	2.38;	Score 41.6;	DB 2;	Length 1597;
Best Local Similarity	44.4%;	Pred. No. 0.57;		

Matches 167; Conservative 0; Mismatches 209; Indels 0; Gaps 0;

```
QY 675 ggcctccccctgtgcacgccaacatgactgtgtcactgacatcaatcgtatgct 734
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 878 GGCATGGCAGGCGGCATCACTACTACCGCTTCTGGTGGCTTCCCTTCCCATGTG 937
QY 735 gctgctgtggtgtcctctctggtggtggtggtggtggtggtggtggtggtg 794
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 938 CCTGCTGCTGGCGCTCTCAACAGGCGATCTGCGCGCGCGGCGGACCGACCA 997
QY 795 gctgaagatggtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 854
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 998 GAAGAGCCGCAAGACAGATCCAGCGGCTGTCTGCTAGCACCGGTCATCTTCTG 1057
QY 855 gtgagtcgtatgtatctactcagcaacaagaacagcaaacagctccacactg 914
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1058 CTGCTTCTGCTGCTACACAGCTGTCTGCTGTGCGCAGCTGTGGAGGCGAGT 1117
QY 915 caagctgtgacatgcccctgcgcgcgaatgcctgtggtggtgtgtgtgtgtgt 974
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1118 CTTGCGCCAAAGGCGCTTTTCACAGCGCTTACACATCTTCTCTGCTCACAG 1177
QY 975 cgaagctcccaagtgacacagtcacagccagagcaagtcacacagggagacgt 1034
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1178 CGTGGCCGACCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1237
QY 1035 caccgagggcgtgagc 1050
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1238 CTTCCGCGGCGGCGCTGC 1253
```

RESULT 10
US-08-687-289A-1
: Sequence 1, Application US/08687289A
: Patent No. 5981195

GENERAL INFORMATION:
: APPLICANT: Fuller, Forrest H.
: APPLICANT: Krapcho, Karen J.
: APPLICANT: Hammerland, Lance G.
: TITLE OF INVENTION: CHIMERIC RECEPTORS AND METHODS FOR
: TITLE OF INVENTION: IDENTIFYING COMPOUNDS AND METHODS AT
: TITLE OF INVENTION: METABOTROPIC GLUTAMATE RECEPTORS AND
: TITLE OF INVENTION: THE USE OF SUCH COMPOUNDS IN THE
: TITLE OF INVENTION: TREATMENT OF NEUROLOGICAL DISORDERS
: NUMBER OF SEQUENCES: 8
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Lyon & Lyon
: STREET: 633 West Fifth Street
: STREET: Suite 4700
: CITY: Los Angeles
: STATE: California
: COUNTRY: U.S.A.
: ZIP: 90071-2066
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
: MEDIUM TYPE: storage
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: IBM P.C. DOS 5.0
: SOFTWARE: FASTSEQ for Windows 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/687,289A
: FILING DATE: July 25, 1996
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 60/001,526
: FILING DATE: July 26, 1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Wauburg, Richard J.
: REGISTRATION NUMBER: 32,327
: REFERENCE/DOCKET NUMBER: 220/004
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (213) 489-1600

TELEFAX: (213) 955-0440
: TELE: 67-3510
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3384 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
US-08-687-289A-1

Query Match 2.28; Score 40.8; DB 2; Length 3384;
Best Local Similarity 46.28; Pred. No. 1.1;
Matches 135; Conservative 0; Mismatches 157; Indels 0; Gaps 0;

```
QY 329 aacgagacgtctgtggagccaggtatctctctctctgtggagacgtgtgtgtg 388
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2091 ACCGCACTGTGGGGGCTCAACCTGCAAGTTCGTGCTGTGTTTCTCTGCACT 2150
QY 389 tctgttgcctgtgtgtgtgaagccagactctcactgtgactctcgcgtctctct 448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2151 AGATTGTATCTGTGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2210
QY 449 ttgggttctgttgcacatcgtctctctctgtgtgtgtgtgtgtgtgtgtgtgt 508
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2211 AGCTGGAAGATGATCATCTTTCATACGTGTCACAGAGGCGCTCTCATGCG 2270
QY 509 tctgtgcccgaagaagacacagcgcccggtgtgtgtgtgtgtgtgtgtgtgt 568
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2271 TCTGATGCGCTACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2330
QY 569 tgaccctgtgtaggtatcatcatatagagtggtgtgtgtgtgtgtgtgtgtgt 620
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2331 GGAAGCTGCGCGAGAACATTCAATGAAGCCAAAGTTCATCAGCTTCAAGTCT 2382
```

RESULT 11
US-08-485-588-3
: Sequence 3, Application US/08485588
: Patent No. 5688938

GENERAL INFORMATION:
: APPLICANT: Edward M. Brown
: APPLICANT: Steven C. Hebert
: APPLICANT: Forrest H. Fuller
: APPLICANT: James E. Garrett, Jr.
: TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
: TITLE OF INVENTION: MOLECULES
: NUMBER OF SEQUENCES: 20
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Lyon & Lyon
: STREET: First Interstate World Center
: STREET: Suite 4700
: STREET: 633 West Fifth Street
: CITY: Los Angeles
: STATE: California
: COUNTRY: USA
: ZIP: 90071
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: FASTSEQ
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/485,588
: FILING DATE: 7 June, 1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: PRIOR APPLICATION DATA: including application
: PRIOR APPLICATION DATA: described below: 9
: APPLICATION NUMBER: 08/353,784
: FILING DATE: 9 December, 1994
: APPLICATION NUMBER: PCT/US/94/12117
: FILING DATE: 21 October, 1994

```

: APPLICATION NUMBER: U.S. 08/292,827
: FILING DATE: 23 August, 1994
: APPLICATION NUMBER: U.S. 08/141,248
: FILING DATE: 22 October, 1993
: APPLICATION NUMBER: U.S. 08/009,389
: FILING DATE: 23 February, 1993
: APPLICATION NUMBER: U.S. 08/017,127
: FILING DATE: 12 February, 1993
: APPLICATION NUMBER: U.S. 07/934,161
: FILING DATE: 21 August, 1992
: APPLICATION NUMBER: U.S. 07/834,044
: FILING DATE: 11 February, 1992
: APPLICATION NUMBER: U.S. 07/749,451
: FILING DATE: 23 August, 1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Heber, Sheldon O.
: REGISTRATION NUMBER: 38,179
: REFERENCE/DOCKET NUMBER: 213/005
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (213) 489-1600
: TELEFAX: (213) 955-0440
: TELEX: 67-3510
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3809 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA to mRNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 373..3606
: OTHER INFORMATION:
:
: US-08-485-588-3

```

```

Query Match      2.2% Score 40.8; DB 1; Length 3809;
Best Local Similarity 46.2%; Pred. No.1.2; Indels 0; Gaps 0;
Matches 135; Conservative 0; Mismatches 157;
QY 329 aacgagacgtctgagaccaggtatctctcttccttgaggacccttgagcctctctgac 388
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2516 ACCGCAAGTGTTGGGGCTCAACCTGCAGTTCTGCTGTTCTCTGCACCTTCATGCG 2575
QY 389 tcgtgttgcctgtgtgtgaagccgaactctccacccttgctctcgagcgtctctct 448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2576 AGATTGTCATCTGTGTATCTGCTCTACACCGCGCCGCCCTCAAGCTACCGCAACGAG 2635
QY 449 ttgggtctctgtcgcacatctgctctctctgtctgagcgctcagctcttgccctcaact 508
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2636 AGCTGGAGGATGAGATATCTTTCATCAGTCGTCACGAGGAGCTCCCTCATGCGCTTGGGCT 2695
QY 509 tccctgcccgaagaacacagggcccgaggctgagtgatcttcaactgtgctctgctgc 568
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2696 TCCTGATCGGCTACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2755
QY 569 tgaccctgtgagaggtcattcaatacagagtgctgatacattcaacctgtgt 620
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2756 GGAAGCTGCCGAGAACTTCATGTAAGCCAAAGTTTCATCACCCTTCAGATGCT 2807

```

```

RESULT 12
US-08-484-565-3
: Sequence 3, Application US/08484565
: Patent No. 5763569
: GENERAL INFORMATION:
: APPLICANT: Edward M. Brown
: APPLICANT: Steven C. Hebert
: APPLICANT: James E. Garrett, Jr.
: TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
: NUMBER OF SEQUENCES: 20
: CORRESPONDENCE ADDRESS:

```

```

: ADDRESSEE: Lyon & Lyon
: STREET: First Interstate World Center
: STREET: Suite 4700
: STREET: 633 West Fifth Street
: CITY: Los Angeles
: STATE: California
: COUNTRY: USA
: ZIP: 90071
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: FASTSEQ
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/484,565
: FILING DATE: 7 June, 1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: PRIOR APPLICATION DATA: including application
: PRIOR APPLICATION DATA: described below: 9
: APPLICATION NUMBER: 08/353,784
: FILING DATE: 9 December, 1994
: APPLICATION NUMBER: PCT/US/94/12117
: FILING DATE: 21 October, 1994
: APPLICATION NUMBER: U.S. 08/292,827
: FILING DATE: 23 August, 1994
: APPLICATION NUMBER: U.S. 08/141,248
: FILING DATE: 22 October, 1993
: APPLICATION NUMBER: U.S. 08/009,389
: FILING DATE: 23 February, 1993
: APPLICATION NUMBER: U.S. 08/017,127
: FILING DATE: 12 February, 1993
: APPLICATION NUMBER: U.S. 07/934,161
: FILING DATE: 21 August, 1992
: APPLICATION NUMBER: U.S. 07/834,044
: FILING DATE: 11 February, 1992
: APPLICATION NUMBER: U.S. 07/749,451
: FILING DATE: 23 August, 1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Heber, Sheldon O.
: REGISTRATION NUMBER: 38,179
: REFERENCE/DOCKET NUMBER: 213/006
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (213) 489-1600
: TELEFAX: (213) 955-0440
: TELEX: 67-3510
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3809 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA to mRNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 373..3606
: OTHER INFORMATION:
:
: US-08-484-565-3

```

```

Query Match      2.2% Score 40.8; DB 1; Length 3809;
Best Local Similarity 46.2%; Pred. No.1.2; Indels 0; Gaps 0;
Matches 135; Conservative 0; Mismatches 157;
QY 329 aacgagacgtctgagaccaggtatctctcttccttgaggacccttgagcctctctgac 388
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2516 ACCGCAAGTGTTGGGGCTCAACCTGCAGTTCTGCTGTTCTCTGCACCTTCATGCG 2575
QY 389 tcgtgttgcctgtgtgtgaagccgaactctccacccttgctctcgagcgtctctct 448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2576 AGATTGTCATCTGTGTATCTGCTCTACACCGCGCCGCCCTCAAGCTACCGCAACGAG 2635
QY 449 ttgggtctctgtcgcacatctgctctctctgtctgagcgctcagctcttgccctcaact 508

```

Db 2636 AGCTGAGAGATGATCATCTTCATCACGTCCACGAGGGCTCCCTCATGTGCCCCGTGCGCT 2695
Qy 509 tccctgcccgaagaacacgagggccgggctgggtgagcttcacatctgctctgcgcgc 568
Db 2696 TCCTGATGCGCTACACCTGCTGCTGCTGCCATCTGCTTCTTCTTCCCTTCAAGTCCC 2755
Qy 569 tgacctgtagaggtacatcaatacagagtggtgctgataccacctggt 620
Db 2756 GGAAGCTCCGAGAACTTCATGAGCCAGTTTCATCACCCTTCAAGATGCT 2807

RESULT 13
US-08-480-751-3
Sequence 3, Application US/08480751
Patent No. 5858684
GENERAL INFORMATION:
APPLICANT: Edward F. Nemeth
APPLICANT: Edward M. Brown
APPLICANT: Steven C. Hebert
APPLICANT: Forrest H. Fuller
APPLICANT: James E. Garrett, Jr.
TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,751
FILING DATE: 7 June, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below: 9
APPLICATION NUMBER: 08/353,784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Hebert, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 213/004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 3809 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 373..3606
OTHER INFORMATION:
US-08-480-751-3

Query Match 2.2%, Score 40.8; DB 2; Length 3809;
Best Local Similarity 46.2%; Pred. No. 1.2;
Matches 135; Conservative 0; Mismatches 157; Indels 0; Gaps 0;

Qy 329 aacgagacctgtgaggaccaggtatcttcctctgaggaaacctggacctctctgc 388
Db 2516 ACCGCACTGGTGGGGGCTCAACCTGCAGTTCCTGTTTCCCTGACACTTCATGC 2575
Qy 389 tcgtgtgacctgtgtgtgtaagccgacctctccacctgtgacctctgagacctctct 448
Db 2576 AGATTGATCTGTGTGATCTGTGCTCTACACCGCGCCCCCTCAAGCTACCGCAACGAG 2635
Qy 449 ttgggtctctgtcgacatctctctctgtctgtgctgacagctctgtgacctctgaact 508
Db 2636 AGCTGAGAGATGAGATCATCTTCATCACGTGCCACGAGGGCTCCCTCATGTGCGGCT 2695
Qy 509 tccctgcccgaagaacacgagggccgggctgggtgagcttcacatctgctctgcgcgc 568
Db 2696 TCCTGATGCGCTACACCTGCTGCTGCTGCCATCTGCTTCTTCTTCCCTTCAAGTCCC 2755
Qy 569 tgacctgtagaggtacatcaatacagagtggtgctgataccacctggt 620
Db 2756 GGAAGCTCCGAGAACTTCATGAGCCAGTTTCATCACCCTTCAAGATGCT 2807

RESULT 14
US-08-943-986-3
Sequence 3, Application US/08943986
Patent No. 5962314
GENERAL INFORMATION:
APPLICANT: Edward M. Brown
APPLICANT: Steven C. Hebert
APPLICANT: James E. Garrett, Jr.
TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,986
FILING DATE: 03-OCT-1997
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/484,565
FILING DATE: 7-June-1995
APPLICATION NUMBER: 08/353,784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117

```

1 TITLE OF INVENTION: MOLECULES
2 NUMBER OF SEQUENCES: 20
3 CORRESPONDENCE ADDRESS:
4 ADDRESSEE: Lyon & Lyon
5 STREET: First Interstate World Center
6 STREET: Suite 4700
7 STREET: 633 West Fifth Street
8 CITY: Los Angeles
9 STATE: California
10 COUNTRY: USA
11 ZIP: 90071
12 COMPUTER READABLE FORM:
13 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
14 COMPUTER: IBM PC compatible
15 OPERATING SYSTEM: PC-DOS/MS-DOS
16 SOFTWARE: FASTSEQ
17 CURRENT APPLICATION DATA:
18 APPLICATION NUMBER: US/08/353,784
19 FILING DATE: 9 December, 1994
20 CLASSIFICATION: 514
21 PRIOR APPLICATION DATA:
22 PRIOR APPLICATION DATA: including application
23 PRIOR APPLICATION DATA: described below: 8
24 APPLICATION NUMBER: PCT/US/94/12117
25 FILING DATE: 21 October, 1994
26 APPLICATION NUMBER: U.S. 08/292,827
27 FILING DATE: 23 August, 1994
28 APPLICATION NUMBER: U.S. 08/141,248
29 FILING DATE: 22 October, 1993
30 APPLICATION NUMBER: U.S. 08/009,389
31 FILING DATE: 23 February, 1993
32 APPLICATION NUMBER: U.S. 08/017,127
33 FILING DATE: 12 February, 1993
34 APPLICATION NUMBER: U.S. 07/934,161
35 FILING DATE: 21 August, 1992
36 APPLICATION NUMBER: U.S. 07/834,044
37 FILING DATE: 11 February, 1992
38 APPLICATION NUMBER: U.S. 07/749,451
39 FILING DATE: 23 August, 1991
40 ATTORNEY/AGENT INFORMATION:
41 NAME: Heber, Sheldon O.
42 REGISTRATION NUMBER: 38,179
43 REFERENCE/DOCKET NUMBER: 209/069
44 TELECOMMUNICATION INFORMATION:
45 TELEPHONE: (213) 489-1600
46 TELEFAX: (213) 955-0440
47 TELEX: 67-3510
48 INFORMATION FOR SEQ ID NO: 3:
49 SEQUENCE CHARACTERISTICS:
50 LENGTH: 3809 base pairs
51 TYPE: nucleic acid
52 STRANDEDNESS: single
53 TOPOLOGY: linear
54 MOLECULE TYPE: cDNA to mRNA
55 FEATURE:
56 NAME/KEY: CDS
57 LOCATION: 373..3606
58 OTHER INFORMATION:
59 US-08-353-784-3
60
61 Query Match 2.2%; Score 40.8; DB 3; Length 3809;
62 Best Local Similarity 46.2%; Pred.No.1.2;
63 Matches 135; Conservative 0; Mismatches 157; Indels 0; Gaps 0.
64
65 329 aacgagaccctcgtggagaccagatattcttcctctcgtggagaccctgggcctctctgc 388
66 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
67 Db 2516 ACCGCAAGGCGGGGGCGTCAACCTGCAGATTCTCTGCTGTTTCTCTGACCTTCATC 2575
68 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
69 389 tcgtcttcgcctcgtgttggtgaagccgcgaattccacactgtgcctctcgtgcgctctct 448
70 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
71 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
72 Db 2576 AGATTGTCAATCTGTGTGATCTGCTCTACACCGCGCCCGCCCTCAAGACTACCGCAACG 2635

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2002, 14:38:27 : Search time 3328.72 Seconds
(without alignments)
11435.438 Million cell updates/sec

Title: US-09-895-686-7
Perfect score: 1819
Sequence: 1 cgcctcgagccctaccacgc.....cttataactcttaaaaa 1819

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vl:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.or:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sts:*
28: em.un:*
29: em.vl:*
30: em.htg.hum:*
31: em.htg_inv:*
32: em.htg_other:*
33: em.htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No. Query Match Length DB ID Description

1	1797	98.8	1866	6	AX083423	AX083423 Sequence
2	1786.8	98.2	1824	6	AF207989	AF207989 Homo sapi
3	1785	98.1	1808	6	AX136163	AX136163 Sequence
4	1753	96.4	1824	9	BC016860	BC016860 Homo sapi
5	1630.4	89.6	1790	6	AX014744	AX014744 Sequence
6	1357.4	74.6	1428	9	HS276102	AJ276102 Homo sapi
7	1324.4	72.8	1326	6	AX014742	AX014742 Sequence
8	1323	72.7	1323	6	AX083413	AX083413 Sequence
9	1088	59.8	192087	2	AC079325	AC079325 Homo sapi
10	1088	59.8	197283	2	AC055863	AC055863 Homo sapi
11	826.8	45.5	191360	2	AL663079	AL663079 Mus muscu
12	826.8	45.5	196975	2	AL669969	AL669969 Mus muscu
13	679	37.3	1662	9	BC004925	BC004925 Homo sapi
14	575.8	31.7	613	6	AX136494	AX136494 Sequence
15	378	20.8	1845	6	AK000249	AK000249 Homo sapi
16	331	18.2	400	11	G14600	G14600 SHGC-11589
17	280.6	15.4	333	6	AX335205	AX335205 Sequence
18	270.6	14.9	77713	9	HUAC004131	AC004131 Homo sapi
19	270.6	14.9	134640	9	AC027130	AC027130 Homo sapi
20	270.2	14.9	1212	6	AX015106	AX015106 Sequence
21	270.2	14.9	1327	9	HS276101	AJ276101 Homo sapi
22	270.2	14.9	2870	9	AF181862	AF181862 Homo sapi
23	270.2	14.9	3852	6	AX015108	AX015108 Sequence
24	268.6	14.8	1402	9	AF202640	AF202640 Homo sapi
25	228.4	12.6	251	6	AX136641	AX136641 Sequence
26	224.4	12.3	2899	10	BC020004	BC020004 Mus muscu
27	223.4	12.3	127219	2	AC023806	AC023806 Mus muscu
28	222.8	12.2	4488	10	AF378831	AF378831 Mus muscu
29	202.4	11.1	161577	9	AC007688	AC007688 Homo sapi
30	202.4	11.1	204272	2	AC007215	AC007215 Homo sapi
31	201.4	11.1	1619	6	AX078375	AX078375 Sequence
32	201.4	11.1	2302	9	AF095448	AF095448 Homo sapi
33	201.4	11.1	2305	9	BC003665	BC003665 Homo sapi
34	201.4	11.1	2446	9	AK001761	AK001761 Homo sapi
35	201.4	11.1	3371	6	AX188348	AX188348 Sequence
36	191.6	10.5	2484	6	AR129188	AR129188 Sequence
37	158.6	8.7	604	6	AX014842	AX014842 Sequence
38	144.4	7.9	1324	10	AF218809	AF218809 Mus muscu
39	140.6	7.7	519	6	AX079652	AX079652 Sequence
40	134.2	7.4	972	6	AX077889	AX077889 Sequence
41	132.6	7.3	1038	9	AF209923	AF209923 Homo sapi
42	132.6	7.3	1089	6	AX207667	AX207667 Sequence
43	132.6	7.3	161577	9	AC007688	AC007688 Homo sapi
44	117.2	6.4	2145	10	AB030197	AB030197 Mus muscu
45	99.4	5.5	47857	2	AC110297	AC110297 Homo sapi

ALIGNMENTS

RESULT 1
LOCUS AX083423 1866 bp DNA
DEFINITION Sequence 115 from Patent WO0112660.
ACCESSION AX083423
VERSION AX083423.1 GI:13185264
KEYWORDS
SOURCE
ORGANISM human.
Human sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1866)
Kato,S. and Kimura,T.
AUTHORS Human proteins having hydrophobic domains and dnas encoding these
proteins
JOURNAL Patent: WO 0112660-A 115 22-FEB-2001;
SAGAMI CHEMICAL RESEARCH CENTER (JP) ; Protegene Inc. (JP)
FEATURES
source location/Qualifiers
1..1866
/organism="Homo sapiens"
/db_xref="taxon:9606"
142..1467
/note="unnamed protein product"

```
/codon_start=1
/protein_id="CAC33306.1"
/db_xref="GI:13185265"
/translation="MAIHKALVMCLGPLELPGANAGHVPSPGSGNLINLYNICD
RSGAKGIVLEAVAGAGIVTFEVLITILVSLPVPQDTRKSLGTOYFLITGLGFLC
LVACVYKPDFSTCASRRLFGLVLAIFCECLAHAFALFKRNKNGPRGWIIFVA
LLLTJLEVIINTMLITTLVRSGESEGPQNSAGNAVASPCA1AMNDFYMALIYVWL
LLLGAFAGAMPALCGRYKRWKRKHGVFLITLTATSAIVWVIMYMYTGNOKHNSPTMD
DPTLALALANMAAEVLEFYIPEVSQVTKSGPSQGDYKPTRGVYETILKEOKGO
SMFEKNKAFSEMPVAAKRPVSPSGNGLSLSVOPTMALMHKVPSCAYDILLP
RATANSQVMSASSTILRAEDMYSASQSHQATTPPKDKNSVFRNPYVMD"
```

BASE COUNT 334 a 602 c 551 g 379 t

ORIGIN

Query Match 98.8%; Score 1797; DB 6; Length 1866;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1797; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

10 cccctaccagcggaaagtaacgaatcgctcagcttgaggagaccacaagcagctggc 69
|||||
70 CCCCTACACGCCGGAAGTAGAGATCGGCTCAGCTGGAGGACCAACGAGACCTGGC 129
|||||
70 cttggagccaggaatgagcattccacaagaagccttggatgctgagctgagctctctc 129
|||||
130 CTGGAGACCAAGATGGCATCCAAAGCCTTGGATGTGCTGGGACTGCTCTCTTC 189
|||||
130 ctgttcccaaggagcttgagcccaaggccaatgtcccaaccgagctgagccaagagctcaac 189
|||||
190 CTGTTTCCAGGGGCTGGGCGCAAGGCCATGTCCACCGGCTTCAGGCAAGGGCTTCAC 249
|||||
190 cccctgtactacaacatgctgagcagctctggaggctgaggacatcgtctcaggagcgctg 249
|||||
250 CCCCTGTACTACAACTGTGTGACCGGCTGGGGCGTGGGCAATGCTGTGAGGCGCTG 309
|||||
250 gctggagcggagcatgtgcaacagcttctgctcaacatcctctgtgagcagctctccc 309
|||||
310 GCTGGGGGGGCAATGTGTACACACCTTGTGCTACCATCTGTGGGCCACGCTCCCC 369
|||||
310 ttgtgagagacacaagaagaagagagctgtgaggagccagagatctctctcttggagg 369
|||||
370 TTTGTGAGGAGACACAAAGAAAGGAGCCTGCTGGGAGCCAGATATCTTCTTGAGGG 429
|||||
370 acccttgagcctctctcgtctgctgctgtgctgctgtgagagccagactctccacgtg 429
|||||
430 ACCCTGGGCGCTTCTGCTGCTGTGGCTGTGGTGGAGGCCGACTTCCACCTGT 489
|||||
430 gctctcggagctctctctcttggagctctgtctgacatctgctctctctctgagcgct 489
|||||
490 GCTCTCTGGGCGCTTCTTGGGGTCTGTGTCGCATCTGCTTCTTGTCTGGCGCT 549
|||||
490 cagctcttggcctcaactctctgagccggagaagaaccaagggagctggagctggc 549
|||||
550 CAGCTCTTGGCTTCACTTCTGTGGCCGGAGAACACAGGGGCCCGGGCTGGGTGATC 609
|||||
550 ttcactgtgctctgctgctgtaaggtlcatlcatcaatagagtgagctgagc 609
|||||
610 TTCACTGTGGCTGTGCTGTACCTGTGAGAGTCAATCAATACAGATGGCTGATC 669
|||||
610 atcaacctggtctggagcagtgagagggcgccctcaaggcaacagcagcgagctgg 669
|||||
670 ATCAACCTGTTGGGGAGTGGGAGGCGGCGCTCAGGCAACAGAGCGAGGCTGG 729
|||||
670 ggcgtgagctctccctctgacatgcaacaatgagacttgcagagcaatcatlaagctc 729
|||||
730 GCGCTGGCGCTCCCTGTGGCATGCGCAACATGAGCTTGTATGAGCATATCTACCTC 789
|||||
730 atgctgagctgctgaggtgacctctcctgaggagcgcccgccctgtgtgagcgctacaag 789
|||||
790 ATGCTGCTGCTGCTGTGGGCTTCTGTGGGCGCTGCGCCGCTGTGTGGCGCTACAG 849
|||||
790 cgcgtgagtaagcattggagcttctgctctcctcaacagcagcactcgttgcataatg 849
|||||

Db 850 CGTGGCGTAGCATGGGGTCTTGTGCTCTCTCACACAGCCACCTCCGTTGCCATATGC 909

Qy 850 gtgggtgagatcgctcatgatacttaagcgaacaagaagcacaacagctcccaactggagat 909
|||||

Db 910 GTGGTGTGGATCGTCAATGTATCTTACGGCAACAAGCAGCAACAGTCCCACTGGGAT 969

Qy 910 gaccccaagctgagcattcgccctcgcgcgcaatgcttggagctctctctcttaagctc 969
|||||

Db 970 GACCCCAAGCTGGGCAATCGCCCTCGCCGCAATGCTGTGGGCTTGTCTTCTTACGTC 1029

Qy 970 atccccgaggtctcccaaggtgaccaaagctccagcccaagaagctccagggagagatg 1029
|||||

Db 1030 ATCCCGAGAGTCTCCAGGTGACCAAGTCCAGCCCAAGCAAGCTACCAAGGGGACATG 1089

Qy 1030 taaccacaaccgggagctggagctatgagacatctcctgaagaagcagaaggtcaagaatg 1089
|||||

Db 1090 TACCCCAAGCGGGCGTGGCTTTCATGATGAGACCATCTGAAAGAGCAAGAGGTCAAGCATG 1149

Qy 1090 ttcgtgagaaacaagccttccatgagatgagccggttgcagctaaagagcggtgca 1149
|||||

Db 1150 TTGCTGGAACAAGGCGCTTTCATGATGAGCGGTTGCAAGAGCGCGTGTCA 1209

Qy 1150 ccatcaagcgggtacaatgagcagctgctgaacaggtgtgtacagcccaatgagatggc 1209
|||||

Db 1210 CCATACAGCGGGTCAATGGGCACTGCTGACAGTGTACAGCCCACTGAGATGGC 1269

Qy 1210 ctgagtcaaaagttccgctccgaagagctatgacatcctcccaaggcgccagcgc 1269
|||||

Db 1270 CTGATGCACAAAGTTCGCTCCGAAGAGACTTACACATCATCTTCCACGGGCGCACCGCC 1329

Qy 1270 aacagcaggtgagatgagcagtgccaactcgacccctgcygagctgaagaatgtaactcgcc 1329
|||||

Db 1330 AACAGCCAGGTGATGGGAGTGCACACCTGCGGGCTGAAGACATGACTCGGCC 1389

Qy 1330 cagaagcaaccagcgagcgaacccgcggaagaagcgcaagaactctcaggtcttgaagaac 1389
|||||

Db 1390 CAGAGCCACAGGGGCGCACCGCCGAAAGACGGAAGACTCTCAAGTCTTTAAGAAC 1449

Qy 1390 cctcaagctgagagctgagctgagcaggtgagcagagagggcggtctgagatcttggagagggc 1449
|||||

Db 1450 CCTACGTGTGGACTGATGATGAGCGGTGGAGAGAGCGGTGGANTTGGGAGGGGCC 1509

Qy 1450 ctgagagacttgcccgagcaaggagactctcaggtctctctccctctgagagccag 1509
|||||

Db 1510 CTGAGAGACTGGCCCGGAGGAAGGAGACTTCAAGGCTCTCTCCCTGGCAGGCGCAG 1569

Qy 1510 caaatgtgcccagaatgtgaaagagcctccctctctgagcaggttggaggtgtgcat 1569
|||||

Db 1570 CAACATGTGCCCGAGATGTGAAGGGCTCCCTCTGTGCCAGTGTGGGTGGTGTAT 1629

Qy 1570 ggtgtccccaaccacactcctcagtgcttgtgagtgagagagcaaccacagcctctgc 1629
|||||

Db 1630 GGGTGTCCCAACCACTCTCTAGTGTGTGTGAGATCGAGAGGCCAACCCAGCTCTGCG 1689

Qy 1630 caagatcaactcggcggatcacaactcagcaaatagtgctcgggltgltgagctgggca 1689
|||||

Db 1690 CAGGATACACTCGGCGGTACACTTCAGCCAAATAGTGTCTCGGGGTGGTGGCTGGGCA 1749

Qy 1690 ggcgtatgttctctctgagatctcctgcaactcaagaagatctccagggcgctcaagcct 1749
|||||

Db 1750 GCGGCTATGTTTCTGTGAGATTCCTGCMACTCAAGAGATTCACAGGCCCTCAGGCT 1809

Qy 1750 ggaatctgactcctctgagagaaagaggtgctcaataaatacatctctgcttcatc 1806
|||||

Db 1810 GGATCTTGCTCTCTGTGAGAGCAAGGTGCTTAATAATACATTTCTGCTTAT 1866

RESULT 2
AF207989 1824 bp mRNA linear PRI 09-MAY-2001
LOCUS AF207989
DEFINITION Homo sapiens orphan G-protein coupled receptor (GPR5C) mRNA,
complete cds.
ACCESSION AF207989

VERSION	AF207989.1	GI:8118031
KEYWORDS	human.	
SOURCE	Homo sapiens	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE	1 (bases 1 to 1824)	
AUTHORS	Brauner-Osborne,H., Jensen,A.A., Sheppard,P.O., Brodin,B., Krogsgaard-Larsen,P. and O'Hara,P.	
TITLE	Cloning and characterization of a human orphan family C G-protein coupled receptor GPRC5(1)	
JOURNAL	Biochim. Biophys. Acta 1518 (3), 237-248 (2001)	
MEDLINE	21210966	
PubMed	11311935	
REFERENCE	2 (bases 1 to 1824)	
AUTHORS	Brauner-Osborne,H., Sheppard,P.O. and O'Hara,P.J.	
TITLE	Direct Submission	
JOURNAL	Submitted (23-NOV-1999) Department of Medicinal Chemistry, The Royal Danish School of Pharmacy, 2 Universitetsparken, Copenhagen 2100, Denmark	
FEATURES	Location/Qualifiers	
source	1..1824	
gene	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	1..1824	
CDS	/gene="GPRC5C"	
	83..1408	
	/note="similar to RAIG1, GPRC5B and GPRC5D"	
	/codon_start=1	
	/product="orphan G-protein coupled receptor"	
	/protein_id="AAF2870.1"	
	/db_xref="GI:8118032"	
	/translation="MAIHKALVMCLGLPLFLPGMAQGHVPPGCSQGLNPLYNYLDCDSFGKGVILAEVAVAGATVTFVFLIIIVASIPFQODKKRSLLGTQVFLIGLGLFCRSVGVKVPDEFTSCARRFLFVGFALFISCLAHVAFNLAKRNGPQGVFTVAALLTLFVITLITLIVAGSGEGDQGNSSAGMAVSPCALANNDPMALILYVLLILGALFLGAMPALCGRRKRRKHKEVPLITFATSVAILVWVYIYTGNGQNSPTVDDPTAIALAANAAAFVLPYIPEVSQYTKSSPEDISQGDMPTRGVEETILKEOKGSMFEENKAFSMDPEVAAKRPVSPYSYNGOLITSVQPTMALMKVPSGCAVDIILPRATNSOVMSANSTLRAEDMYSASOSHQAATPPKDGKNSQVFRNPYVD"	
BASE COUNT	337 a	579 c 535 g 373 t
ORIGIN		
Query Match	98.2%	Score 1786.8; DB 9; Length 1824;
Best Local Similarity	99.6%	Pred. No. 0;
Matches 1802; Conservative	0;	Mismatches 7; Indels 1; Gaps 1;
QY	10	ccctaccagccggaagatagcagatcgcgctgaagcttgtaggagaccacaacagagctggc 69
DB	11	ccctaccagccggaagatagcagatcgcgctgaagcttgtaggagaccacaacagagctggc 70
QY	70	ctggagaccaggaatgagccatccacaagaagccttgatgtagtgcgtgtagagctgcctcttc 129
DB	71	ctggagaccaggaatgagccatccacaagaagccttgatgtagtgcgtgtagagctgcctcttc 130
QY	130	ctgttcccaaggggcttggtcccaagggccatgtccaccgccgagctgcagccaagagctcaac 189
DB	131	ctgttcccaaggggcttggtcccaagggccatgtccaccgccgagctgcagccaagagctcaac 190
QY	190	cccctgtactacaacctgtgaacgcgctgagggcgctgagggagcatgctctgaagacgtg 249
DB	191	cccctgtactacaacctgtgaacgcgctgagggcgctgagggagcatgctctgaagacgtg 250
QY	250	gctggggcgggcatgtgtaaccaagcttgctgtcacaacatcaatccttggttgcaagcctccc 309
DB	251	gctggggcgggcatgtgtaaccaagcttggttgctcacaacatcaatccttggttgcaagcctccc 310
QY	310	tttgagcaggaacacaagaagagcgctgctggtgagaccagagattcttctcttgagg 369
DB	311	tttgagcaggaacacaagaagagcgctggtggtgagaccagagattcttctcttgagg 370

QY	370	acccttgagcctctctctgacctcggttgacctgtgtgtgtaagcccgacctctccacctgt	423
Db	371	acccttgagcctctctctgacctcggttgacctgtgtgtgtaagcccgacctctccacctgt	430
QY	430	gacctctgagcctctctctgacctcggttgacctgtgtgtgtaagcccgacctctccacctgt	489
Db	431	gacctctgagcctctctctgacctcggttgacctgtgtgtgtaagcccgacctctccacctgt	490
QY	490	caactcttgacctcaactctctctgacctcggttgacctgtgtgtgtaagcccgacctctccacctgt	549
Db	491	caactcttgacctcaactctctctgacctcggttgacctgtgtgtgtaagcccgacctctccacctgt	550
QY	550	ctcaacttgacctctctctgacctcggttgacctgtgtgtgtaagcccgacctctccacctgt	609
Db	551	ctcaacttgacctctctctgacctcggttgacctgtgtgtgtaagcccgacctctccacctgt	610
QY	610	atacaacttgacctctctctgacctcggttgacctgtgtgtgtaagcccgacctctccacctgt	669
Db	611	atacaacttgacctctctctgacctcggttgacctgtgtgtgtaagcccgacctctccacctgt	670
QY	670	gaccttgacctctctctgacctcggttgacctgtgtgtgtaagcccgacctctccacctgt	729
Db	671	gaccttgacctctctctgacctcggttgacctgtgtgtgtaagcccgacctctccacctgt	730
QY	730	atgctgctgacctctctctgacctcggttgacctgtgtgtgtaagcccgacctctccacctgt	789
Db	731	atgctgctgacctctctctgacctcggttgacctgtgtgtgtaagcccgacctctccacctgt	790
QY	790	cgctgagctgaagcctctctgacctcggttgacctgtgtgtgtaagcccgacctctccacctgt	849
Db	791	cgctgagctgaagcctctctgacctcggttgacctgtgtgtgtaagcccgacctctccacctgt	850
QY	850	gtgtgtgacctctctctgacctcggttgacctgtgtgtgtaagcccgacctctccacctgt	909
Db	851	gtgtgtgacctctctctgacctcggttgacctgtgtgtgtaagcccgacctctccacctgt	910
QY	910	gaacccacgctgacctctctgacctcggttgacctgtgtgtgtaagcccgacctctccacctgt	969
Db	911	gaacccacgctgacctctctgacctcggttgacctgtgtgtgtaagcccgacctctccacctgt	970
QY	970	atccccgagctctctctgacctcggttgacctgtgtgtgtaagcccgacctctccacctgt	1029
Db	971	atccccgagctctctctgacctcggttgacctgtgtgtgtaagcccgacctctccacctgt	1030
QY	1030	taacccacgctgacctctctgacctcggttgacctgtgtgtgtaagcccgacctctccacctgt	1089
Db	1031	taacccacgctgacctctctgacctcggttgacctgtgtgtgtaagcccgacctctccacctgt	1090
QY	1090	ctctgtgagaaacagcctctctgacctcggttgacctgtgtgtgtaagcccgacctctccacctgt	1149
Db	1091	ctctgtgagaaacagcctctctgacctcggttgacctgtgtgtgtaagcccgacctctccacctgt	1150
QY	1150	ccataacagctgacctctctgacctcggttgacctgtgtgtgtaagcccgacctctccacctgt	1209
Db	1151	ccataacagctgacctctctgacctcggttgacctgtgtgtgtaagcccgacctctccacctgt	1210
QY	1210	ctgtgtgacaaagtctctctgacctcggttgacctgtgtgtgtaagcccgacctctccacctgt	1269
Db	1211	ctgtgtgacaaagtctctctgacctcggttgacctgtgtgtgtaagcccgacctctccacctgt	1270
QY	1270	aacagcagctgtgacctctctgacctcggttgacctgtgtgtgtaagcccgacctctccacctgt	1329
Db	1271	aacagcagctgtgacctctctgacctcggttgacctgtgtgtgtaagcccgacctctccacctgt	1330
QY	1330	cagagcacaacagctgacctctctgacctcggttgacctgtgtgtgtaagcccgacctctccacctgt	1389
Db	1331	cagagcacaacagctgacctctctgacctcggttgacctgtgtgtgtaagcccgacctctccacctgt	1390
QY	1390	ccctacagctgtgacctctctgacctcggttgacctgtgtgtgtaagcccgacctctccacctgt	1449
Db	1391	ccctacagctgtgacctctctgacctcggttgacctgtgtgtgtaagcccgacctctccacctgt	1450
QY	1450	ctgagagactgtgacctctctgacctcggttgacctgtgtgtgtaagcccgacctctccacctgt	1509

|||||
Db 1451 CTGAGACCTGGCCCCGGGCAAGGAGCTCTCAGGCTCCTCCCTCCCTGGCAGG-CCAG 1509
1510 caaatgtgccccagatgtggaagagcctccctctctgccaagtgttggtggtgtgcat 1559
1510 |||||||
Db 1510 CAACATGTGCCCCAGATCTGGAAAGGCTCCCTCTCTCCAGTGTGGGTGTAT 1569
1570 ggggtcccccaactcctcaagtgtgtgtgagtcgaagagcaacccagcctctgc 1629
1570 |||||||
Db 1570 GGGGTGCCCAACCTCCTAGTGTGTGTGTGAGTCAAGAGCCACCCAGCCTCCTGC 1629
1630 caggatacctcggcgtcacactccagcaaatagtgtctcggggtgtgtgtgga 1689
1630 |||||||
Db 1630 CAGGATACCTCGCGGCTCACACTCCAGCCAAATAGTCTTCGCGGAGTGTGCTGGCA 1689
1690 ggcgctgttctctgtgagatctctgcaactcaagagactccagcgctcagacct 1749
1690 |||||||
Db 1690 GCGCTATGTTCTCTGGAGATCTCTGCAACCTCAAGAGACTCCAGGCTCAGGCT 1749
1750 ggaactgtcctctgtgagaacaagggtgcctaataacatctctgtcttaac 1809
1750 |||||||
Db 1750 GGATCTGTCTCTCTGTGAGSAGACAGGCTCCTAATAATACATTTCTCTTATTAA 1809
1810 tcttaaaaa 1819
1810 |||||||
Db 1810 AAAAAAAAA 1819
RESULT 3
AX136163 1808 bp DNA linear PAT 30-MAY-2001
LOCUS AX136163
DEFINITION Sequence 85 from Patent EP1067182.
ACCESSION AX136163
VERSION AX136163.1 GI:14272571
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1808)
AUTHORS Ota,T., Isogai,T., Nishikawa,T., Kawai,Y., Sugiyama,T. and
Hayashi,K.
TITLE Secretory protein or membrane protein
JOURNAL Patent: EP 1067182-A 85 10-JAN-2001;
Helix Research Institute (JP)
FEATURES
source location/Qualifiers
1..1808
CDS
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="unnamed protein product"
/codon_start=1
/protein_id="CA39724.1"
/db_xref="GI:14272572"
/translation="MAIHKALVNCGLPLFLFPGAMAQGHVPPGCSOGLNPLLYNLCD
RSCAMGVLEAVAGAGIVTFLTLILVSLPVDPTKRSLLFLARKNHGPGVIFVLA
LVFACVYKPEPSTCAARRFLFGLFAIFCFSCLAARVLAFLARKNHGPGVIFVLA
LITIVELIITFEMLIITTLIRVSGEGGPOGSSAGAAVVAAPLANNDFMALIIVML
LLGALFGLAMPALCGRIKRRKRRKRPVLELTATSVAILVIMYITVYTKGNQHSPTWD
DPTLALALANAAAFVLEFVYIPEVSOTKSSPQOSYOGDMYPTRGVYETILKEQGO
SMEVENKAFSMDPVAARVPVSGYNGQLTSVYQPTSMALMHKVPSPGAYDILLP
RATANSQVMSANSTLRAEDMYSQASHOATPPKDKNSGVFRNPYVMD"
BASE COUNT 321 a 580 c 531 g 376 t
ORIGIN

Query Match 98.1%; Score 1785; DB 6; Length 1808;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1796; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 10 ccctccaccagccggaagctacagatcggtcagctcgtgagagaccacaacagacctgac 69
11 |||||||
12 |||||||
Db 13 CCTTACCACGCCGGAAGTACGAGTGGCTCAGCTGAGGAGACCAACCAAGACCTGGCC 72

QY 70 ctggagaccagatggtccatccacaagccttggatgtgtccttggactgctctctc 129
71 |||||||
Db 73 CTGGGAGCCAGAGATGGGCATCCACAAAGCCTTGATGTGCTCCGGAGCTCCTCTTC 132
130 ctgttcccaagggtccttggccaaggccaatgtcccaaccggtcgaagccaagcctaac 189
130 |||||||
Db 133 CTGTCCAGAGGCGCTGGGCCAGAGGCCATGTCCACCCGCGCTCAGACCAAGGCTTCAAC 192
190 cccctgtactaacaacccgtgtgagcgtcttggagcgttggagatcgtccctgggaagccgtg 249
190 |||||||
Db 193 CCGCTGACTACAACTGTGTGACCGCTGTGGGCGTGGGCAATGCTGTGGAGGCCGTG 252
250 gcttggcgggcatgttcacacagcttgtctcaacatcatctccttggcgaagcctccc 309
250 |||||||
Db 253 GCTGGGCGGGCATTTGTCACACACTTGTGTGTACATCATCTGTTGGGAGGCTCC 312
310 ttgttgcaagacacacaagaaaggagcctgtcgtgggaagccaagatctctctcttgggg 369
310 |||||||
Db 313 TTGTGCAAGGACACCAAGAAAGGAGCTGTGGGAGCCAGATATCTTCTTCTGGGG 372
370 accctggacctctctctcgtgtgttggcgtgtgtgtggaagccgaactctccacctgt 429
370 |||||||
Db 373 ACCTGGGCGCTTCTCTGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 432
430 gacctcggcgtctcctctcttggagttctgttcgcaatctgtctctcttctgtcgcgct 489
430 |||||||
Db 433 GCCTCTGGCGCCTTCTCTTGTGGGTTCTGTGTGCCATCTGCTTCTGTGTGTGGCGCT 492
490 caagctcttgcctcaactctcctgtgcggaagaacaacagggcccggtgtgtgtgtc 549
490 |||||||
Db 493 CACCTTTTGGCCCTCAACTTCTGCGCCGGAAGAACCAAGGCGGCCCGCGGCTGTGATTC 552
550 ttaactgtgctctgtcgtctgacccctgtgagagttcatcatcaataagatggctgctc 609
550 |||||||
Db 553 TTCACTGTGCTCTGTCTGTGACCTCTGTGAGGTATCATCATATACAGATGCTATTC 612
610 atcaccttgtctcgggagctgtgcgaagcgccctcaaggccaacagcaagcgaagctgtg 669
610 |||||||
Db 613 ATCACCGCTGTTGGGGGAGTGGGAGGCGGCCCTCAGGGCAACACAGCGAGCGCTGG 672
670 ggcgtgacctccctctgtgtccatcgcaacaatgtgaattgtcatgtcaactcatcgtc 729
670 |||||||
Db 673 GCGCTGGCCTCCCGCTGTCGATGCGCAACATGTGACCTTTGATGAGCACTATCTACGTC 732
730 atgtctgtcgtcgtgtgtgtcctctctgtgggcccgtggccgcccgtgtgtgcccgtacaag 789
730 |||||||
Db 733 ATGCTGCTGCTGCTGGGTGCTCTCTGTGGGCGCTGGCCGCGCTGTGTGGCGCTTACAG 792
790 cgtctgctgaagcaatgtgtgtcttctgtctcctcaacaaagcactcctgttgcataatg 849
790 |||||||
Db 793 CGCTGGGGTAAAGCATGGGGGTCTTGTGTCTCTTCACACAGACCATCTCGTTGCCATATGG 852
850 gtgggtgtgatcgtcatgtatacttaactgagcaagaagcaagcaacagttcccaactgtgat 909
850 |||||||
Db 853 GTGTGTGTGATCGTCAATGTACTTACGGCAACAGCAGCAGCAACAGTCCCACTCTGGAT 912
910 gaccaccaagctgggacatcgccctcgcgccaatgtccctgggacctctgtcctcttcaagtc 969
910 |||||||
Db 913 GACCCACGCGGGCCATCGCCCTGCGCCGCAATGCTGGGCGCTTGTGTCTTCTTACGTC 972
970 atccccgaagctctccaggtgacccaagtccagcccaagcaagaactcaccaggggagatg 1029
970 |||||||
Db 973 ATCCCGAGGTCTCCAGAGTACCAATCCAGCCCAAGCAAGCAACTACAGGGGAGATG 1032
1030 taccaccaaccggggtgtgggtatgtgacatctctgaagaagcaagaaggttcaagatg 1089
1030 |||||||
Db 1033 TACCCACCGCGGGCGTGGCTATGAGACCATCTGAAAGAGCAGAGGGTCAAGAGCATG 1092
1090 ttgttgagaagaagcctcttccatgtgatagtgagccgttgcagcttaagaagccgggttca 1149
1090 |||||||
Db 1093 TTCTGTGAGAAACAAGGCGCTTTTCCATGATGAGCGGCTTCCAGCTAAGAGCGCGGTCTCA 1152

Contact: (Dickson, Mark) mdg@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers
R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 22 Row: h Column: 1.

FEATURES

Location/Qualifiers

BASE COUNT	381 a	558 c	517 g	368 t
ORIGIN				

Query Match	96.4%	Score 1753	DB 9	Length 1824
Best Local Similarity	99.7%	Pred. No. 0		
Matches 1756	Conservative	0	Mismatches 5	Indels 0
				Gaps 0

QY	59	cagaagccctggagccctggagagccagaatgagccatccacaagaacccctgtgtgattgctctggagac	118
Db	1	CAGAGCCCTGGCTGGGAGCCAGAGATGGCCATTCACAAAGCCTTGGTGATGTGCTGGGAC	60
QY	119	tgacctctctccctgtgtccccaaggagcctgggcccaggagcatgtgccaccggcctgcaagcc	178
Db	61	TGCTCTCTCTCTGTTCTTCCAGAGGGGCGCTGGGCCACAGGGCCATGTCCACCCGGCTGGAGCC	120
QY	179	aaggcctcaaacccctctgactacaacctgtgtgacgcctctvgggcgctggggagcatctgc	238
Db	121	AAGGCTCACAACCCCGCTGATCTAACCTGTGTAGCCGCTCTGGGGGCTGGGGCATCTGCC	180
QY	239	tggagagccttgcctggggagggagcatgtgtaacagctttgtgtcaacatcalccctgttg	298
Db	181	TGGAGGCGCTGGCTTGGGGGGGCGGCGCATTTGTACACAGCTTTTGCTTCACATCATCTGTGTG	240
QY	299	ccaagcctccctctgtgtgcagagacaccagaagaaacggagcctgtctggggagcccaagtaact	358
Db	241	CCAGCCTCCCTTTGTGTGCAGAGACACCAAGAAAGGAGCCTGCTGGGGAGCCACAGTATCT	300
QY	359	tcctctctggggagaccttgggacctctctctgcctcgtgttgcctgctgtgtgtggaagccgact	418
Db	301	TCCCTCTGGGGAGCCCTGGGCGCTTCTTGCTCTGTGTGTTCCTGCTGTGGTGAAGCCCGACT	360
QY	419	tctcaacctgtgacctctggagccttcacctctctctcttctggggattctgtgtgcacatcttctct	478
Db	361	TCTCCACCTGTGTCTCTGTGGCGCTTCTCTTGTGGGGTGTGTGTGGCATCTGTCTTCTT	420
QY	479	gtcttgagcgctcaagctctcttgccctcaactctccttgcccggagaagaacacagggccccggt	538
Db	421	GCTGTGGCGGCTACGCTTGTGCCCTCAACTTCTTGCCCGGGAAGAACCGGGGCCCGGG	480
QY	539	gctgggtgatacttcaactgtgactctgtcctgtgtaacctgtgtagaggtcatcatcaatcaaac	598

[illegible]

QY	1679	gagcgtcgggcaagcgcctatgcttctcttggagatctccgtcaactcaagaagcttccagg	1738
Db	1621	GGGGCTGGGACACGGCTAAGTTTCTGTGAGATTCTTCGAACCTCAAGAGACTTCCAGG	1680
QY	1739	cgcctagccttgatcttctcctctgttgaagaacaaggtgcctcaataatacttct	1798
Db	1681	CGCTCAGCGCTTGACATCTTCTCTCTCTGTGAGAACAGAGGTCCCTAATATACATTCT	1740
QY	1799	gctttatttaactcttaaaaa	1819
Db	1741	GCTTATTATTAACAAAAA	1761
RESULT	5		
AX014744			
LOCUS	AX014744	1790 bp	DNA
DEFINITION	Sequence 3 from Patent WO9953054.		Linear
ACCESSION	AX014744		
VERSION	AX014744.1	GI:10041010	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 1790)		
TITLE	Michalovich,D., Hill,J., Medhurst,A. and Pangalos,M.		
JOURNAL	Axot4 g-protein-coupled receptor		
	Patent: WO 9953054-A 3 21-Oct-1999;		
FEATURES	SMITHKLINE BECHAM PLC (GB)		
source	Location/Qualifiers		
	1..1790		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
BASE COUNT	324 a 560 c 522 g 379 t	5 others	
ORIGIN			
Query Match	89.6%; Score 1630.4; DB 6; Length 1790;		
Best Local Similarity	97.7%; Pred. No. 4.4e-297;		
Matches 1758: Conservative	0; Mismatches 26; Indels 15; Gaps 10;		
QY	13	tcaccagccggaagtacagatcgcgtcagcctcgtgagaggaaccacaagcctgcgtg	72
Db	1	TCACCAACCCGGAATAGACAGTGGGCTCAGCCTGGAGGAGACCAACAGAGCTGGCTG	60
QY	73	ggagccaggaatggcatccacaagaagccttgtatgtcgtcgtgagcgtcctctctctg	132
Db	61	GGAGCCAGATTGTGGCATTCACCAAGCCCTGGTGTGATGTGCTGGACTGCTCTCTCTG	120
QY	133	ttcccaagaggtcctgggcccagagcagatgtcccaaccgcgcctgtgacgaagcctaaacc	192
Db	121	TTTCCCAAGGGGCTTGG--CAGGCAATGTCTCCACCCGGCTGCAAG--CAAGGCTCAACCC	177
QY	193	ctgtactaaacactgtgtgacacgctccttgagggtgttgatcgtctcgttgaggcgttg	252
Db	178	CTGTACTAACACCTGTGTGACCGCTGTGGGGGCTGGGAGCATGTCTGTGAGGCGCTGCT	237
QY	253	ggggcggcgatgtgtccacagcttgtgtgtcaacatcatctcgtgtggcagcgtccctt	312
Db	238	GGGGGGGATTTGTACCACTGTTGTGTGCTCACACATCATCTGTGTGGCAGGCTCCCTTT	297
QY	313	gttcaggaacacaaagaagagcgtgtcgtggagaccaagatacttccctcttgaggac	372
Db	298	GTTCAGGACACCAAGAAAGGAGCTGCTGTGGGACCCAGATTTCTTCTTGTGGGTAC	357
QY	373	ctggagcctctctgcctcgttgttgcgtgtgtgtgaagcccgacttccactctgacc	432
Db	358	CTGGGCTCTTCTGTGCTGCTGTTGTTGCTGTGTGTTGTTGTTGTTGTTGTTGTTGTTG	414
QY	433	tctcggcgtctcctcttggggatcttgttgcacatcgtctctctgtctggtggctaac	492
Db	415	TTCTGGCGCTTCTTGTGGGGTGTGTGTGGCATCTCTTCTTGTGTGTGGCGCTTAC	474

	BASE COUNT	324 a	560 c	522 g	379 t	5 others
ORIGIN						
Query Match	89.6%	Score 1630.4	DB 6	Length 1790		
Best Local Similarity	97.7%	Pred. No. 4.4e-297				
Matches 1758	Conservative	0	Mismatches 26	Indels 15	Gaps 10	
QY	13	tcaccagccggaagtacagatcggctcaagccttggagaggaaccaacagagccttgcc	72			
Db	1	TCACCAGCCGGAAGTACAGATCGGCTCAGCCTGGAGGGAGCCACAGAGCGTGGCTG	60			
QY	73	ggagccaggaatggccatccacaagccttgtatgtctccttggagctcctcttcc	132			
Db	61	GGAGCCAGATTGGCATTCACCAAAACCTTGGTGAATGGCTGGGACTGCTCTTCTCG	120			
QY	133	ttcccgagagccttggcccgagagcagatgtcccaaccggccttggacgaagagcctaaacc	192			
Db	121	TTCCCGAGGGGCTTGGG--CAGGGCAATGTCCACCCGGCTGGAG--CAAGGCTCAACCCC	177			
QY	193	ctgtactaaacactgtgtgacacgctccttgggagcttgggacatcgtcttggagagccgttgc	252			
Db	178	CTGTACTAACACCTGTGTGATCCGCTGTGGGGGCTGGGAGCATGTCTGGAGCCGTGGCT	237			
QY	253	ggggcgggcattgttccacacgtttgtgtctcaacatcatcccttgggtggcagctccccc	312			
Db	238	GGGGCGGGCATTTGTACCACTGTTTGTGCTCACATCATCTGTGTGGCCAGCCTCCCTTT	297			
QY	313	gtcagagacaccaaagaagagcctgtccttgggagccagatacttccctcttgggagc	372			
Db	298	GTTCAGAGACCAACAAGAGGAGCCTGCTGGGGACCCAGATATTCTTCTTGGGTAC	357			
QY	373	ctggagccttcttgcctcgttgttgcctgtgtgtgaaagccggaacttccacactgtgc	432			
Db	358	CTGGGCGCTCTTGCCTGCTGTTTGTGCTGTGTGGTGAAGCCGCACTT--TCTGTGTCC	414			
QY	433	tctcggagcttccctcttggggatcttgttcgcaatcgtctctcttgccttggcggtcac	492			
Db	415	TTCTGGCGCTTCTTTTGGGGTGTGTGTGGCATCTCTTCTTGTGTGGCGCTTAC	474			

QY	493	gtctctgcccctcaactctctctgcccggaaagcaacacggygcccgggctctggtgatcttc	552
Db	475	gtcttttgcctcaacttctctgcccggaaagcaacacggygcccgggctctggtgatcttc	534
QY	553	actgtgctctctgtctgtgaacctgtgtagaagtcatacacaataagaatgtgcatactc	612
Db	535	actgttgctctgtgtgctgacccctgttagagggatcatcatcatatcagatgtgctatcatc	594
QY	613	aacctgtgtcggggcaagtgtgcagagggcgagccctcaaggccaagaagcgaggtctgggccc	672
Db	595	accctgtgttgggggcagtgagtgccagagggccgcccctcagagccacagcagcgttgggccc	654
QY	673	gttgagcctccccctgttgccatctgcgcaacatgtgacttgcataatgacatactaatgtatg	732
Db	655	gtggcgctccccctgttgccatctgcgcaacatgtgacttgcataatgacatactaatgtatg	714
QY	733	ctgtctgtctgtggtgtgcctctctctggtgggctctgagccgcgcctgtgtgtgacgtactaaagcgc	792
Db	715	ctgtctgtctgtggtgtgcctctctctggtgggctctgagccgcgcctgtgtgtgacgtactaaagcgc	774
QY	793	tggggtgaagaatgaggggttcctttgtgcccctccacaagaagcactccgtgttgacataatggtg	852
Db	775	tggggtgaagaatgaggggttcctttgtgcccctccacaagaagcactccgtgttgacataatggtg	834
QY	853	gtgtgtgaatcgtcatgtactacttcagcgcaacaaagcagcacaagtcctccactctgtgatgac	912
Db	835	gtgtgtgaatcgtcatgtactacttcagcgcaacaaagcagcacaagtcctccactctgtgatgac	894
QY	913	cccaagctgtgacatcgccctcgccgccaatgtcctctgggccccttcgtctctctactagatcac	972
Db	895	cccaagctgtgacatcgccctcgccgccaatgtcctctgggccccttcgtctctctactagatcac	954
QY	973	cccgaggtctccccaaggtgtgaaccaagttcccaagcgaagcaagttaccaggggagacatgtac	1032
Db	955	cccgaggtctctccccaaggtgtgaaccaagttcccaagcgaagcaagttaccaggggagacatgtac	1010
QY	1033	cccaaccgagggg-gtgtgacatgtagaaccaatctctgaagaagcagaaggtctgaagacatgt	1091
Db	1011	cccaaccgagggg-cgtggcgtatgtagaaccaatctctgaagaagcagaaggtctgaagacatgt	1070
QY	1092	cgtgtgagaaacaaggtccttttccaatgtagagccggtgtgagctaaagagccggtgtcaccc	1151
Db	1071	cgtgtgagaaacaaggtccttttccaatgtagagccggtgtgagctaaagagccggtgtcaccc	1130
QY	1151	ataaagcggtgtacaatgtggtgaagctgtctgtgacaggtgtgtacagagcccaatgtagaatgcccct	1211
Db	1131	ataaagcggtgtacaatgtggtgaagctgtctgtgacaggtgtgtgtacagagcccaatgtagaatgcccct	1190
QY	1212	gattgacaagaatttcgcgtccgaagaaggtcttaacgacatcatctcccagaagggcacgcgcca	1271
Db	1191	gattgacaagaatttcgcgtccgaagaaggtcttaacgacatcatctcccagaagggcacgcgcca	1250
QY	1272	caagccaggtgtatgtggcaggtgtccaactcgagccctgtgcggctgtgaagacatgtatctcgcccca	1331
Db	1251	caagccaggtgtatgtggcaggtgtccaactcgagccctgtgcggctgtgaagacatgtatctcgcccca	1309
QY	1332	gagccacaagcggtgcgcacacgcgc-gcaaaagacggtcagaactctagaatctttagaaacc	1390
Db	1310	gagccacaagcggtgcgcacacgcgc-gcaaaagacggtcagaactctagaatctttagaaacc	1369
QY	1391	cgtaagctgtgtggaactgtatgaagctgtggtgaggaagaagaggtgcgaatttggggagggccc	1450
Db	1370	cgtaagctgtgtggaactgtatgaagctgtggtgaggaagaagaggtgcgaatttggggagggccc	1429
QY	1451	tgaaggaactgtgcccgggccaaggaagatctt-ccaggtcctctctccccccttgccagccag	1509
Db	1430	tgaaggaactgtgcccgggccaaggaagatctt-ccaggtcctctctccccccttgccagccag	1489
QY	1510	caacatgtgtcccacagatgtgtgaaaggtcctccctctctgcagatgttvtgtgtgtgtcat	1569
Db	1490	caacatgtgtcccacagatgtgtgaaaggtcctccctctctgcagatgttvtgtgtgtgtcat	1549

Oy	1570	ggggtctcccaaccactctctcaagtggttggtgagtcgagggccaaccccaagctctctgc	1629
Db	1550	tgggtgtccccaaccactctctcaagtggttggtgagtcgagggccaaccccaagctctctgc	1609
Oy	1630	cagatcaaccctcgagggcgacactcccaagcaaatggtcttcgggggtggtgagtcgagga	1689
Db	1610	cagatcaaccctcgagggcgacactcccaagcaaatggtcttcgggggtggtgagtcgagga	1668
Oy	1690	ggcctatggtctctcttgagagatccctcgacaactccaagagctcccaagcgctcaagcct	1749
Db	1669	ggcgctatggttctcttgagagatccctcgacaactccaagagctcccaagcgctcaagcct	1728
Oy	1750	ggatctgtctctctctgtgagagacaagggtgcctaataatacatcttcgtctaataa	1808
Db	1729	ggatctgtctctctctgtgagagacaagggtgcctaataataatacatcttcgtctaataa	1787
RESULT	6		
LOCUS	HSA276102	1428 bp	Linear
DEFINITION	Homo sapiens mRNA for GPRC5C protein.		PRI 30-Jul-2000
ACCESSION	AJ276102		
VERSION	AJ276102.1	GI:9588668	
KEYWORDS	gprc5c gene.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	Robbins,M.J., Michalovich,D., Hill,J., Calver,A.R., Medhurst,A.D., Gloger,I., Sims,M.A., Middlemiss,D.N. and Pangalos,M.N.		
TITLE	Molecular cloning and characterisation of two novel retinolic acid-inducible orphan G-protein-coupled receptors (GPRC5B and GPRC5C)		
JOURNAL	Genomics 76, 8-18 (2000)		
REFERENCE	2 (bases 1 to 1428)		
AUTHORS	Michalovich,D.		
JOURNAL	Direct Submission		
FEATURES	Submitted (22-MAR-2000) Michalovich D., Bioinformatics, SmtHkline Beecham, Third Avenue, Harlow, Essex CM19 5AW, UK		
Source	1. 1428		
gene	/organism="Homo sapiens"		
CDS	/db_xref="taxon:9606"		
	103..1428		
	/gene="gprc5c"		
	103..1428		
	/gene="gprc5c"		
	/codon_start=1		
	/product="GPRC5C"		
	/protein_id="CAC00633.1"		
	/db_xref="GI:9588669"		
	/translation="MAITHKALVMLGLPLFLFPRANAQGHVPPCSOGILNPYINLCD		
	RSQMGVILEAVVAGAVITVFVLTITLIVASLPVQDRKSLSTOVFLIGTGLDFCC		
	LVFCAVCAVPDEFTSCASRFLFGLVFALFCISGLAHVAFLNARNGPRQVLFVAA		
	LLLELLEVITINTEWLITLLVSGSGEGPGNSSGMAVASPCAANNDPFWALLIYVIM		
	LLLAFGAPALPCGRKRMKRGVFPVLTATSVAILVMVYITVGNKOHNSPTMDND		
	DPFLAIALANMAAFVLEFVYIPEVSOVTKSSPEQSGODMYPRFGVGYETLLKEQKQ		
	SMFPEKNKFSMDEPVAARPRSPISGTINGOLITSVYPTKEALMHNKPSSEAYIILIP		
	RATNSQVMSGSANSTFLREDYSAQSHQDAATPPDRGNNOVFRNPYVMD"		
BASE COUNT	254 a 463 c 421 g 290 t		
ORIGIN			
Query Match	74.6%;	Score 1357.4;	DB 9;
Best Local Similarity	99.9%;	Pred. No. 1.1e-245;	Length 1428;
Matches 1358;	Conservative 0;	Mismatches 1;	Indels 0;
		Gaps 0;	
Oy	49	gggaccacaacccagagcctcgagccttgagagccaagatcgccatctccacaagaacgcttgtaag	108
Db	70	gggaccacaacccagagcctcgagccttgagagccaagatcgccatctccacaagaacgcttgtaag	129
Oy	109	tgccttgagactgcctctctctctgtctcccaaggagccttgagcccaaggacatgtcccaacc	168

|||||
Db 130 TGCCTGGGAGTCCCTCTCTCTGTTCCAGGGGCTGGGCCAAGGGCCATGTCCACACC 189
Oy 169 ggctbcaagcaaggcctaaccocccctbactaacaacctgtgtaccgctctgggagcgtg 228
Db 190 GGCTGACACCAAGGGCTCAACCCCTGTACTACAACTGTGTGACCGCTGTGGGGCTGGG 249
Oy 229 ggcacgcctcgtgagagccgtgtgtgtggcgagcattgtccacaagtgtgtgtccacatc 288
Db 250 GGCAATCGTCTGGAGAGCGGTGGCTGGGGCGGCAATGTACACAGCTTGTGTCTACCACTC 309
Oy 289 atcctgtgtgccaagcctcccttgtgtgagaacacaaagaaacggagcctgtgtggagac 348
Db 310 ATCTGTGTGGCGAGCGCTCCCTTGTGTGACAGAACCAAGAAACGAGCGCTGTGGGGAC 369
Oy 349 caggtatctctctctgtggagacctgtgacctctctctcgtgtgtgtgtgtgtgtgt 408
Db 370 CAGGATATCTTCTCTTGTGGGACCTTGGGCCCTTCTCTGCTGTGTGTGTGTGTGTGTG 429
Oy 409 aagcccgactctccacctgtgacctcgtcgtcgagcctccctctgtgggtgtgtgtgtgt 468
Db 430 AAGCCCGACTTCTCCACTGT 489
Oy 469 tgcctctctgt 528
Db 490 TGCCTCTCTGT 549
Oy 529 gggcccgagggt 588
Db 550 GGGCCCGAGGGGT 609
Oy 589 atcaatacagaagt 648
Db 610 ATCAATACAGAGT 669
Oy 649 ggcacaagcagaagcgt 708
Db 670 GGCACACACGACGCGAGGT 729
Oy 709 gtcatgtgacacatctactgt 768
Db 730 GTCATGTGACACTATCTATCTATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 789
Oy 769 gccctgt 828
Db 790 GCCCTGT 849
Oy 829 ggcacctcgt 888
Db 850 GCCACCTCCGTTGCAATGT 909
Oy 889 cacaacagttccacactgt 948
Db 910 CACAACTGT 969
Oy 949 gccctgt 1008
Db 970 GCTTGT 1029
Oy 1009 caaagctcagaaggt 1068
Db 1030 CAAGCTACACGAGGGGAGATGTATACCCACCGGGGGCGGTGTGTGTGTGTGTGTGT 1089
Oy 1069 gagcagaaggt 1128
Db 1090 GAGCAGAAAGGCTCAGAGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1149
Oy 1129 gcaagtaagaagcgt 1188
Db 1150 GCAGCTAAGAGGCGGT 1209
Oy 1189 taccagccactgt 1248

Db 1210 TACAGCCCACTGAGATGGCCCTGTATGACAAAGTTCCTCCGGAAGGACTTACGACATC 1269
Oy 1249 atcctcccaagggtccacacgcgaacagcaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1308
Db 1270 ATCTCTCCACCGGCCACCGCCACACAGCGATGTATGTGAGTGTGCACTTCAACCTGTGG 1329
Oy 1309 gctgaagacatgtactcgtgcccagaagccacagaagcgtgacacacgcgcgaagaagcgaag 1368
Db 1330 GCTGAGACATGTACTGTGGCCAGAGCCACAGAGGGGGCCACACCGCCGAAGAGCGGACAG 1389
Oy 1369 aactcaggtccttagaaccctcagtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1407
Db 1390 AACTCTCAGGTCTTTAGAAACCTTACTGTGTGGACTGA 1428

RESULT 7
AX014742
LOCUS AX014742 1326 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 1 from Patent WO953054.
ACCESSION AX014742
VERSION AX014742.1 GI:10041009
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 1326)
AUTHORS Michalovich D., Hill J., Medhurst A. and Pangalos M.
TITLE Ax04, a protein-coupled receptor
JOURNAL Patent: WO 953054-A 1 21-Oct-1999;
SMITHKLINE BEECHAM PLC (GB)
FEATURES
source 1..1326
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 232 a 431 c 386 g 277 t
ORIGIN

Query Match 72.8%; Score 1324.4; DB 6; Length 1326;
Best Local Similarity 99.9%; Pred. No. 1.8e-239;
Matches 1325; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 82 atggccatccacaagcctgt 141
Db 1 ATGGCATTCACAAAGCCTTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 60
Oy 142 gccctgt 201
Db 61 GCTGTGGGCCAGGGCCATGTCCACCGCGTGCAGGCCAAGGCTTCACACCTGTACTATAC 120
Oy 202 aacctgt 261
Db 121 AACTGT 180
Oy 262 atgttccacaagttgt 321
Db 181 ATTGTACACAGTTGT 240
Oy 322 accaagaagaagcgt 381
Db 241 ACCAAGAAACGAGACCTGT 300
Oy 382 ttctgtcctcgt 441
Db 301 TTTCCTCTCGT 360
Oy 442 ttccctctgt 501
Db 361 TTTCCTCTGT 420
Oy 502 ctcaactctgt 561

Dh 421 CTCACCTTCCTGGCCCGGAAGAACACAGGGCCCGGGGGCTGGGTGATCTTCACTGTGGCT 480
Qy 562 ctgctgtacacctgtgtatagatcatcaataacagatgtgtgtatcatcaccctggtc 621
Dh 481 CTGCTCTGACCTGTGTAGAGTCAATCAATCAAGAGTGGCTGATCATCACCTGTGGT 540
Qy 622 cgggagcagtcgagagggcgccctcagggcaacagcagcagcctgtgggcgtggtcc 681
Dh 541 CGGGGAGTGGCGAAGGGCCCTCAAGGGACACAGCAGCGAGGCTGGGGCTGCCCTCC 600
Qy 682 cccctgtgcatcgcaacatltgtaattgtcaatgtgcaatcattacatgtctgtgtc 741
Dh 601 CCTGTGGCGTCCCAACATGAGACTTGTGATGACACTCATCTACGTCATGCTGCTG 660
Qy 742 ctgggtgtccttctgtggggcgtgcccgcctgtgtgtgcccgtacaaagcgtgctaa 801
Dh 661 CTGGGTGCTCTCTGGGGGCTGGCCGCTGTGTGGCGGTAAAGCCCTGGCGTAAG 720
Qy 802 catgggtcttctgtctcctcaacacagccaactcgtgtgcaatgtgggtgtgtgtatc 861
Dh 721 CATGGGGTCTTTGTGCTCTCAACACAGCCACTCCGTTGCCATGTGGTGTGGATC 780
Qy 862 gtcatgtatactacgcaacaaagcagcaacagtcaccaactgtggatgtgaacccacg 921
Dh 781 GTCATGTATCTTACGGCAACAGACAGACACAGTCCACTGGGATGTAGCCACGCTG 840
Qy 922 gcatgtccctctgcccgaatgtcctgtggcctgtctcttctacgtcatcccgagtc 981
Dh 841 GGCATGCTCCCTCCCGCAATGGCTGGGCTTGTGCTTTCATGCAATGCCGGAAGTC 900
Qy 982 tcccaagtgcaacagtcaccaagcccaagcaagctaccagaagagatgtaccacacg 1041
Dh 901 TCCCAAGTGAACAGTCCAGCCAGACCCAGAACAGTACCAGAGGGACATGTACCCAC 960
Qy 1042 ggcgtgtggtatgtagaacccatccctgaagaagcagaaggttcaagatgttctgtga 1101
Dh 961 GCGGTGGGTATGAGACCATCTGAAAGAGCAGAAAGGTCAGAGCATGTTCGTGGAGAC 1020
Qy 1102 aagagcttctcattgtatgtgagccggtgtgcaagcagaagccggtgttaccataga 1161
Dh 1021 AAGGCTTTTTCATGTGAGTGGCGGTGACGTAAAGGCGGTGTCCATCATACAGG 1080
Qy 1162 taacaatggagagctgtgtagcaagtcgtgtaccagcccaatgtgaatgtgctgtga 1221
Dh 1081 TACAATGGGAGAGCTGCTGACCAAGTGTACCAAGCCCACTAGATGGCTGATGACAAA 1140
Qy 1222 gtctcgtccgaagagctttagcaatcattcctccacagggccacccgcaaacagcag 1281
Dh 1141 GTTCCGTCGAAGAGCTTACGACATCATCTCCACAGGGCCACCCCAACAGCCAG 1200
Qy 1282 atgtggcaggtgtcaaatcgtgaacctgtgagagacatgtactcgtgcccagaagcca 1341
Dh 1201 ATGGGAGTGTCCCAACTGACCTGCGGGCTGAAGACATGTACTCGGCCAGAGCACA 1260
Qy 1342 ggggcaacacgcccgaagaagcagcaagaacttccaagttctttagaacccttactgt 1401
Dh 1261 GCGGCCACACCGCCGAAGACGGCAGACACTCTCAGGTCTTTAGAAACCCCTAC 1320
Qy 1402 gactga 1407
Dh 1321 GACTGA 1326

RESULT 8
AX083413 1323 bp DNA linear PAT 28-FEB-2001
LOCUS AX083413
DEFINITION Sequence 105 from Patent WO0112660.
ACCESSION AX083413
VERSION AX083413.1 GI:13185250
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
1 (bases 1 to 1323)
AUTHORS Kato, S. and Kimura, T.
TITLE Human proteins having hydrophobic domains and dnas encoding these proteins
JOURNAL Patent: WO 0112660-A 105 22-FEB-2001;
SAGAMI CHEMICAL RESEARCH CENTER (JP) ; Proteogene Inc. (JP)
FEATURES
source 1..1323
BASE COUNT 232 a 431 c 384 g 276 t
ORIGIN
Query Match 72.7%; Score 1323; DB 6; Length 1323;
Best Local Similarity 100.0%; Pred. No. 3.2e-239;
Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 82 atggccatcccaaaagcctgtgtatgtgtccttggagactgtcctctcctgtccagag 141
Dh 1 ATGGCCATCCCAAAAGCCTTGTGTATGTGCTGGAGCTGCTCTCTCTGTTCCTCCAGGG 60
Qy 142 ggcctggcccagggccatgtcccaacccggtgtcagccaagcctcaacccctgtactac 201
Dh 61 GCGTGGGCCAGGGCCATGTGCCACCGGGCTGCAGCCAAAGCCCTCAACCCCTGTACTAC 120
Qy 202 aacctgtgtacgcgtctgtggcgctgtgggcatcgtcctgaagcgtgtgctgtggcgagc 261
Dh 121 AACCTGTGTACCGCTCTGGGGCGTGGGCGCATCTGCTGAGAGCGCTGGGCGGGCGGC 180
Qy 262 atgttccacagtttgtgtgtccatcatctcctgtgtgccaagcctccctctgtgtgaag 321
Dh 181 ATGTGTACCAAGTGTGTGTCTCACCATATCTGTGTGCGCAAGCTCCCTTGTGTCAAGAC 240
Qy 322 acaagaagaagagcctgtcgtgggagaccagatctctcctctgtggagacctgtggcctc 381
Dh 241 ACCAAGAAGAGAGCCCTGCTGGGGAGCCAGGATATTCTTCCTTGGGGAGCCCTGGGCTC 300
Qy 382 tctgtcctgtgtgttgcctgtgtgtgtgaagccgacttctcactgtgtcctcctggcgc 441
Dh 301 TTGTGCTGTGTTGCTGTGCTGTGTGAAGCCCACTTCTCACCTGTGCTCTCGGCGCG 360
Qy 442 ttctcttgggtgttctgttcgttcgcatctgtctctgtctgtgtgggtgttaacgtttggc 501
Dh 361 TTCTCTTTGGGGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
Qy 502 ctcaacttccctgtggcggagaaacacagcggcccgaggctgtgtatcttcaactgtgtgct 561
Dh 421 CTCACCTTCTGCGCCGCGGAAGAACACAGGGGCCCGGGGCTGGGTGATCTTACGTGGCT 480
Qy 562 ctgctgtgaacctgtgttagaggtatcatcaataacagatgtgtatcatcaacctgtgtc 621
Dh 481 CTGCTGTGACCCCTGTGTAGAGTATCATCAATCAAGTGGCTGATCATCAACCTGTGTT 540
Qy 622 cgggagcaggtgtgagagggcgccctcaaggtcaacagcagcagcgtgtgggcgtgtcc 681
Dh 541 CGGGGCAAGTGGCGAAGGGCGGCTCAAGGCAACAGCGAGGCTGGGCGTGGCTGC 600
Qy 682 cccctgtgcatcgcaacatgtgaacttgtcatgtgcaatcattactatgtatgtctgtgt 741
Dh 601 CCTGTGTCATGCGCAACATGTGACTTGTGATGTGACACTCATGTACATGCTGCTGCTG 660
Qy 742 ctgggtgtccttctgtgggcgtgtgcccgtgtgtgtgcccgtacaaagcgtgtgctlaag 801
Dh 661 CTGGGTGCTTTCCTGGGGGCTGTGGCCGCTGTGTGGCGCTAAAGCGCTGGCGTAAG 720
Qy 802 catgggtcttctgtgtccttcaacacagccaactcgttgcataatgtgtgtgtgtatc 861
Dh 721 CATGGGTCTTTGTGTCTTCACACAGCAACCTCCGTTGCCATATGAGTGTGTGATC 780
Qy 862 gtcatgtatactacgcaacaaagcagaacagcagtcaccaactgtgtgagccccagcgt 921

	source	1.	.191360 /organism="Mus musculus" /db_xref="taxon:10090" /chromosome="11" /clone="RP23-273C14" /clone_id="RPCI-23"
	misc_feature	1..11633 /note="assembly_fragment:01402 fragment_chain:1 clone_end:SP6 vector_side:left"	
	misc_feature	11734..22769 /note="assembly_fragment:03027 fragment_chain:1"	
	misc_feature	22870..36573 /note="assembly_fragment:01950 fragment_chain:1"	
	misc_feature	36674..46266 /note="assembly_fragment:00462 fragment_chain:1"	
	misc_feature	46367..51724 /note="assembly_fragment:00283 fragment_chain:1"	
	misc_feature	51825..62425 /note="assembly_fragment:02841 fragment_chain:1"	
	misc_feature	62536..99778 /note="assembly_fragment:01029 fragment_chain:1"	
	misc_feature	99879..102684 /note="assembly_fragment:02931 fragment_chain:1"	
	misc_feature	102785..118140 /note="assembly_fragment:00217 fragment_chain:1"	
	misc_feature	118241..126500 /note="assembly_fragment:03406 fragment_chain:1"	
	misc_feature	126601..130506 /note="assembly_fragment:02614 fragment_chain:1"	
	misc_feature	130607..138567 /note="assembly_fragment:02597 fragment_chain:1"	
	misc_feature	138668..141877 /note="assembly_fragment:00946 fragment_chain:1"	
	misc_feature	141978..159504 /note="assembly_fragment:00985 fragment_chain:2"	
	misc_feature	159605..185733 /note="assembly_fragment:02648 fragment_chain:2"	
	misc_feature	185834..188643 /note="assembly_fragment:01377 fragment_chain:3"	
	misc_feature	188744..191360 /note="assembly_fragment:02017 fragment_chain:3"	
BASE COUNT	46220 a 48864 c 47109 g 47544 t 1623 others		
ORIGIN			
Query Match	45.5%:	Score 826.8;	DB 2; Length 191360;
Best Local Similarity	84.0%:	Pred.No.4.7e+146;	
Matches 946;	Conservative	0; Mismatches 177;	Indels 3; Gaps 1;
Oy	9 gccctaccacagccggaagtacgatcgatcgtgcatacggaggaccaaacagaccttg	68	
Dd	9478 GACCTTGAGGAGCAATTCGTCTCCCTCTCGTCTCTGAAGGACCACACTAAAGCCTGG	9419	
OY	69 cctggagaccagatgccatcaccaaaagccttggtagtgttcctggactgctctcttt	128	

Db	9418	CTTG6GAGCCAGATGGCCACTACAGAACTTGCTGATGTGCTTACAGACTGCTTCTT	9339
QY	129	ccgtgtcccaaggagccctgagcccaaggagccatgtcccaaccggctgtgcaagccaaagccctcaa	188
Db	9358	C---TTCCCAAGAGACCTTGGCCCAAGATCATGCCCCACCTGGTTGCACCCAGATCTGGA	9302
QY	189	ccccctgtactaaacctgtgtgacgcgtctctgggagcgttgagatcgtctctgaagccgt	248
Db	9301	TCCCTCTACTACAACTCTGTATACCGCTCGGGGGGCTGGGGCATTTGCTTGGAGGAGCT	9242
QY	249	ggctgggggggggcatgtgacccaacgtttgttcaacatcatccctgfyggccaacccccc	308
Db	9241	GGCTGGAGCAGGCATATACAGCAATTTGTTATACATCATCTTGGTGACCTTCC	9182
QY	309	cctgtgtcaagaaacacaagaacgagagcctgtctgggagaccagatctctctctctcgtgg	368
Db	9181	ATTGTGTGAGGACACTAAGAAAGGGAGCCTCTTGGGGACCCAGTGTTCTTCTGTGGG	9122
QY	369	gaaccctgggacctctctctgacctgtgtgtgacctgtgtgtgaagccgaactctccacctg	428
Db	9121	CACCTGGGTCCTTCTTGCCTCGTGTGTTCCATGCTGTGGTGAAGCCGGCATTTCTACTCG	9082
QY	429	tgcctctcggagctctcctctctttggggttctgtttgcacatctctctctgtctctggggc	488
Db	9061	TGCTCTTCAGAGCTTCCTCTTTTGGGGTCCTGTTTGGCATATCTCTCTCTGTGTGAGC	9002
QY	489	tcaagctcttggccccctcaactctctgccccgaagaacacagggcccgaggctgtgtgat	548
Db	9001	TCAGTCTCTTCCCTCAACTTCTTACCCGGAAGAACATGGGGCCCGCAGGCTGGGTAT	8942
QY	549	cttcaactgtgctctgtctgtgtaacctggtagaagttcaatcaatcaatlaagagttgctgat	608
Db	8941	CTTCAACGTGGCGCTGCTGCTCAACCTTGTGGAGTCAATTAACACCGAGTGGCTTAT	8882
QY	609	catacccttggtcttggggcaggtgtgcaaggcgagccctcaaggcaacagcaagcgcaagctg	668
Db	8881	CATCAACCTGGTATAGGGGAGGTGGCCAGGTACACCCCTGGGCAATGTACATGCCGACTC	8822
QY	669	ggccgttggccctccccctgtgacatgcacaatagatgttgcattgacatgcaactaactgt	728
Db	8821	GACCATGACCTCTCGTGTGTCATCGCCAAAGATGAGACTTGTTCATGTGCTCTCATCTACGT	8762
QY	729	catgtctgtgctgtctgtgagtgacctctctctgggggaccttggccgacctgtgttggcgctaaa	788
Db	8761	AATGCTGTGTGCTGTGACGGCCTTCCTAGAGAGCCTGGCCACCTTGTGTGGCGCTTCAA	8702
QY	789	ggcgtgcgtgaagcatgagggtctttgtgtctctcaacacagccaacctcgtttgcaatg	848
Db	8701	GGCGTGGGGGAAACACCGGGCTCTTGTGTGTGCACCACTGTCAATCTGCATCGGCATCTG	8642
QY	849	ggtgtgtgtgtgtgtatgtataacttaaggcaacaagagagcaacaacttcccaacttggga	908
Db	8641	GGTGTATGTGATTTCTATGTACACTTACGGCAACGACGACACATTAAGCCCACTTGGA	8582
QY	909	tgaaccaagctgtgccaatgcacctgcgcgcgaatgactgtggactctgtctctcttctactgt	968
Db	8581	TGACCCCAACGTTGGCATTGTCCTCGGTGCAATGCTGTGACCTTGTGTCTTCTAATGT	8522
QY	969	catccccaaggtctcccaagttgaccaatgtccagatccagagaacaagatctacaggggagac	1028
Db	8521	CATCCTCTGAGGTCTCACAGGTGACAAACCTTACCCAGAACAGACTCAACAGGGGAGAT	8462
QY	1029	gtacccaaccggggcgltgggctgtatagacatctctgaaagaagcagaaggttcaagaagat	1088
Db	8461	GTACCCGACCCGAGGGGTGGGCTATGTAGACATCTCTGAAGGAGCAGACAGGGCCAGAGCAT	8402
QY	1089	gtctgtgagaaacaaggcctcttccatgtatgtgaagccgagttgagct	1134
Db	8401	GTTTGTGAGAACAAAGCATTTTCTATATGAGCAACACAGCTCAGGT	8356

RESULT 12
AL669969/c

LOCUS	AL669969	196975 bp	DNA	linear	HTG 30-JAN-2002
DEFINITION	Mus musculus chromosome 11 clone RP23-254J18, *** SEQUENCING IN				
ACCESSION	AL669969				
VERSION	AL669969..3	GI:18477062			
KEYWORDS	HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.				
SOURCE	house mouse.				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
TITLE	1 (sites)				
JOURNAL	Plumb.B.				
	Direct Submission				
	Submitted (29-JUN-2002) Wellcome Trust Sanger Institute, Hinxton,				
	Cambridgeshire, CB10 1SA, UK. E-mail enquiries:				
	humgenrysanger.ac.uk Clone requests: clonerequest@sanger.ac.uk				
	On Feb 1, 2002 this sequence version replaced gi:18375930.				
COMMENT	----- Genome Center				
	Center: Wellcome Trust Sanger Institute				
	Center code: SC				
	Web site: http://www.sanger.ac.uk				
	Contact: humgenrysanger.ac.uk				
	----- Project Information				
	Center project name: BM254J18				
	----- Summary Statistics				
	Assembly program: XGAP; version 4.5				
	Sequencing vector: plasmid; 108752; 100% of reads				
	Chemistry: Dye-terminator Big Dye; 77% of reads				
	Chemistry: Dye-terminator; 22% of reads				
	Consensus quality: 193975 bases at least Q40				
	Consensus quality: 194270 bases at least Q30				
	Consensus quality: 194637 bases at least Q20				
	Insert size: 195775; sum-of-constants				
	Insert size: 185198; agarose-fp				
	Quality coverage: 11.57x in Q20 bases; sum-of-constants quality				
	coverage: 12.31x in Q20 bases; agarose-fp				

	* NOTE: This is a 'working draft' sequence.				
	* This record will be updated with the finished sequence				
	* as soon as it is available and the accession number will				
	* be preserved.				
FEATURES					
SOURCE	Location/Qualifiers				
	1..196975				
	/organism="Mus musculus"				
	/db_xref="taxon:10090"				
	/chromosome="11"				
	/clone="RP23-254J18"				
	/clone_lib="RPCR-23"				
	1..5154				
	/note="assembly_fragment:02093				
	fragment_chain:1				
	clone_end:SP6				
	vector_side:left"				
	5255..35776				
	/note="assembly_fragment:01566				
	fragment_chain:1"				
	35877..55304				
	/note="assembly_fragment:01716				
	fragment_chain:1"				
	53405..66379				
	/note="assembly_fragment:00116				
	fragment_chain:2"				
	66480..82119				
	/note="assembly_fragment:02883				
	fragment_chain:2"				
	82220..88071				
	/note="assembly_fragment:03245				
	fragment_chain:2"				
	88172..108526				
	/note="assembly_fragment:00436				
	fragment_chain:2"				
	108627..1115927				
	/note="assembly_fragment:04402				

[illegible]

QY	729	catcctcctgctgctgaggtgctctccctcctgagggagccctggccgcgcgcgttgctgagccgttaacaa	788
Db	130986	AATGCTGTTGCTGCTGACGGCGCTTCCTAGAGACCTGGCGCCACTTGTGTGGCGCCTTCAA	130927
QY	789	gcgcctgcgtlaagcatgagtgctcttgctgctctccaccacagccacctcgtltgacatalg	848
Db	130926	GCGCGGGGGAACACGGGGGCTTGTGTGTCTGACCACTGTCACTTCATTCGACATCTG	130867
QY	849	ggctgctgtgagtcgltcalgtlaactlaactlaccgcaacaagcagacaacagltcccaactlgyga	908
Db	130866	GGTGTATGATGATGTGCATGATACACTAGCGCAACAGACGACCATATACCCCACTGGGA	130807
QY	909	tgaccccaacgctgagccatcgcgccctctgcgcgcgaatgctctgggacctctgctctctlaagt	968
Db	130806	TGACCCCACTGTGGCGCATATGGCCCTCGCGCCAAATCCCTGGACCTTGTGTCTTCTTATGT	130747
QY	969	catccccaagctctcccaagctgtaaccaagtccaagccacagacagaacactcaacagggagacat	1028
Db	130746	CATCCCTGAGGCTCTACAGGTGACCAAACTTAGCCCAACACAGAGCTACAGGGGACAT	130687
QY	1029	gtaccccaaccccgaggcgctlgygctalbgagacatccctgaagaagcagaaaggtlcaagacat	1088
Db	130686	GTACCCGACCCGAGGGGGGGGCTATGAGAACCATCTCTGAAGGAGCGAGACGGGCGCAAGCAT	130627
QY	1089	gttcgtlgygaagaacaagcctttccatctgatalgagagccggtltagact	1134
Db	130626	GTTTGTGGAGAACAGGCAATTTCTATGTGAGCAACACCACTTCATAGGT	130561

RESULT	13
BC004925	
LOCUS	
DEFINITION	BC004925 1662 bp mRNA linear PRI 12-JUL-2001
ACCESSION	
VERSION	
KEYWORDS	homo sapiens, similar to G protein-coupled receptor, family C, group 5, member C, clone MGC:10304 IMAGE:3622005, mRNA, complete cds.
SOURCE	BC004925
ORGANISM	BC004925.1 GI:13436247
REFERENCE	human.
AUTHORS	Homo sapiens
TITLE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
JOURNAL	1 (bases 1 to 1662)
REMARK	Strausberg, R
COMMENT	Direct Submission Submitted (21-MAR-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA NIH-MGC Project URL: http://mgc.nci.nih.gov Contact: MGC help desk Email: cgapbs-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada info@cgsc.bc.ca Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Gun, Letitia Hsiao, Martin Krzywinski, Reta Kutschen, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smalls, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL, Plate: 13 Row: b Column: 23
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7020202.

FEATURES		Location/Qualifiers			
SOURCE		1..1662			
		/organism="Homo sapiens"			
		/db_xref="taxon:9606"			
		/clone="MGC:10304 IMAGE:3622005"			
		/tissue_type="Pancreas", adenocarcinoma"			
		/clone_id="NH_MGC_39"			
		/lab_host="DH0B-R"			
		/note="Vector: pOTB7"			
		647..964			
		/codon_start=1			
		/product="Similar to G protein-coupled receptor, family C,			
		group 5, member C"			
		/protein_id="AAH04925.1"			
		/db_xref="GI:13436248"			
		/translation="MAIYHKAIVMLGLPLFLPQAMAQGHVPRSCSGNLNLYNLCD			
		RSQMGIVLEBAVAGAGIVTFTVLIILIVASLPFYQDTKKRSILSTQLRGCHNHAGTMM			
		GSC"			
BASE COUNT	377 a 493 c 520 g 272 t				
ORIGIN					
Query Match	37.3%, Score 679; DB 9; Length 1662;				
Best Local Similarity	98.6%; Pred. No. 5,1e-118;				
Matches	685; Conservative 0; Mismatches 10; Indels 0; Gaps 0;				
QY	1125	ggttcagactaagagaccggtgtcaccatacagcgggtatcaatlyggcagctgtcagcaag	1184		
Db	909	GGACCCACGCTAAGAGGGCGGTGTACCCATACAGCCGGTACAAATGGCGAGCTGTGACCGAG	968		
QY	1185	tgtttacacagcccaatctagatcttgcccttatgtacacaagaattccgtccgaagagcttaca	1244		
Db	969	TGTGTACCAAGCCCATGTGAGATGTGCCCTTAATGACAAAGTTCGTTCCGAAGAGCTTACGA	1028		
QY	1245	catcaacctccacacgacgacacccgcaacagcagcaggtgatctggagcagttgcacactcgacct	1304		
Db	1029	CATCATCTCTCCACGGGGCCACCGCCACAGCAGCGAGTAMTGGCAGTGCACACTCGACCT	1088		
QY	1305	gcgggcgtgaagacatgtactcggcccaagagccacacagcgcgcacacgcgcgaagaacgg	1364		
Db	1089	GGGGCGTGAAGACATGTACTCGGCCACAGGACACAGGGCGGCACACCGCCGAAAGACGG	1148		
QY	1355	caaggaactctcaagtcctttagaaacaccttaogtctgggaactctgaatcagcgcggtggcgagga	1424		
Db	1149	CAAGAACTCTCAGGCTCTTGAAGAACCCCTACCTGTGGGACTGTGATCAGTCAGCGGTGGAGGA	1208		
QY	1425	gaagcgcgcgaattctgggaagggccctctgaagacctctggcccgcgcaaggaactctccaggt	1484		
Db	1209	GAGCGGGGGCGATTGTGGGGAGGGCCCTTGAGAGACTGTGGCCCGGGCAAGGAGACTTTCACAG	1268		
QY	1485	ctctctctctccctctgtgcagcgcacgaacaatgtgtgcccaagatgttgaaagggcctccctct	1544		
Db	1269	CTCCTCTCTCCCTGTGGCAGGCCACACAACATGTGCCGCCCGAGATCTGAAAGGCGCTCCCTCT	1328		
QY	1545	ctgcagcagttcttggtctggtgtgtcattcagtggtgttcccccaaccctccctcagtggttttggaat	1604		
Db	1329	CTGCCAGTGTGGGTGGGTGTGCATGTGGGTGTCCACACCCATCTCTCAGTGTGGAGAT	1388		
QY	1605	cgaagagccaaccccaagcctctctgcacagatcaactctcggcgggttcaacactccagcaaatat	1664		
Db	1389	CGAGGAGCCCAACCCACAGCTCTTGCCACGAGATCACTCGGGGGTTCACACTCCAGCCAAATA	1448		
QY	1665	gtctctctcgggggttggttgagctgggcagcgcctaagtctctcttgaaatctccttcaacttca	1724		
Db	1449	GTGTTCTCTGGGGTGTGGGTGTGGCAGCGGCTATGTGTTTCTGTGAGATTCCTCAACCTCA	1508		
QY	1725	agagaactcccaagcgcgtcagagccttgatctgtccctctgtgtgaagaaacaaggtgtgctta	1784		
Db	1509	AGAGACTTCCCGAGGGGCTCAGGCTCAGGATCTTGTCTCTGTGAGAGAACAAAGGTGCTTA	1568		
QY	1785	ataaatacattctgctttatataactttaaaaaa	1819		
Db	1569	ATAAATACATTTCTGCTTATTAAAAA	1603		

RESULT 14
LOCUS AX136494 613 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 416 from Patent EP1067182.
ACCESSION AX136494
VERSION AX136494.1 GI:14272898
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthelia; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 613)
AUTHORS Ota,T., Isogal,T., Nishikawa,T., Kawai,Y., Sugiyama,T. and Hayaashi,K.
TITLE Secretory protein or membrane protein
JOURNAL Patent: EP 1067182-A 416 10-JAN-2001;
Helix Research Institute (JP)
FEATURES
source 1. 613
location/Qualifiers
BASE COUNT 90 a 200 c 175 g 143 t 5 others
ORIGIN
Query Match 31.7%: Score 575.8; DB 6; Length 613;
Best Local Similarity 99.1%: Pred. No. 1.5e-98;
Matches 577; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

AK000249
LOCUS AK000249 1845 bp mRNA linear PRI 22-FEB-2000
DEFINITION Homo sapiens cDNA FLJ20242 fls, clone COLF6369.
ACCESSION AK000249
VERSION AK000249.1 GI:7020202
KEYWORDS Oligo capping; fls (full insert sequence).
SOURCE Homo sapiens colon mucosa cDNA to mRNA, clone-1lb:COLF clone:COLF6369.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthelia; Primates; Catarrhini; Homnidae; Homo. 1 (sites)
AUTHORS Melanabe,K., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Nakamura,Y., Isogal,T. and Sugano,S.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished (2000)
REFERENCE 2 (bases 1 to 1845)
AUTHORS Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogal,T., Shibahara,T., Tanaka,T. and Nakamura,Y.
TITLE Direct Submission
JOURNAL Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science, University of Tokyo, Department of Virology, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry for Japan: cDNA full insert sequencing: Research Association for Biotechnology: cDNA library construction, 5'-8' end one pass sequencing: Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).
FEATURES
source 1. 1845
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="COLF6369"
/tissue_type="colon mucosa"
/clone_1lb="ColF"
/note="cloning vector pME18SFJ3"
60..548
/note="unnamed protein product"
/codon_start=1
/protein_id="BA91030.1"
/db_xref="GI:7020203"
/translation="MAIHKALVACLGLPLFPBQSYQGDYPPRGVYETLKEQK QSMFEVKKAFSDPYAKRPVPSYSGNQLTSYQPEMLMHRVPSGAYDIL PRAPFANSQVMSANSTLRADMTLSAQSHQAATPPKDKNSQVSELTQROGANNOYFL RL"
CDS
BASE COUNT 391 a 569 c 519 g 366 t
ORIGIN
Query Match 20.8%: Score 378; DB 9; Length 1845;
Best Local Similarity 100.0%: Pred. No. 2.2e-61;
Matches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```
QY 1242 cgacatcatcctcccaaggccacgccaagccagtgatgagtgagtgccaaatcgac 1301
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 353 CGACATCATCTCTCCCAAGGCCACCGCCAAACAGGTGATGGGCACTGCGAC 412
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1302 cctgcgggctgaagacatgtactcgccagagccacacgagccacacgcgcaaga 1361
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 413 CCGCGGGGTGAAGACATGTACTCGGCCAGAGCCACAGGGCGCCACACCGCGAAGA 472
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1362 cggcaagaactctcaagt 1379
    |||||||||||||||||||
Db 473 CGGCAAGAACTCTCAGGT 490
    |||||||||||||||
```

Search completed: September 22, 2002, 17:16:01
Job time: 9454 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2002, 09:26:51 ; Search time 31.69 seconds
(without alignments)
1545.710 Million cell updates/sec

Title: US-09-895-686-1
Perfect score: 2326
Sequence: 1 MAINKALVCLGLPLFLPPG.....ATPPKDGKNSQVFRNPYWD 441

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802.*
1: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2326	100.0	441	21	AAV57283 Human GPCR protein
2	2326	100.0	441	21	AAV50932 Human fetal brain
3	2326	100.0	441	22	AAV39849 Human polypeptide
4	2326	100.0	441	22	AAB88591 Human hydrophobic
5	2326	100.0	441	22	AAB88359 Human membrane or
6	2326	100.0	486	22	AAU14166 Human novel protei
7	2325	100.0	441	20	AAV49156 G-protein coupled
8	2320	99.7	460	22	AAW41635 Human polypeptide
9	2226	95.7	479	21	AAV44273 Human Metabotropic
10	2027	87.1	400	20	AAV36302 Human secreted pro
11	2004.5	86.2	461	20	AAV49157 G-protein coupled

12	1221	52.5	231	22	AAU14402 Human novel protei
13	1200.5	51.6	296	21	AAV44274 Mouse Metabotropic
14	1020.5	43.9	302	21	AAV50933 Human fetal brain
15	733	31.5	403	20	AAV32141 Human G-protein co
16	733	31.5	403	21	AAB43085 Human GPCR protein
17	733	31.5	403	21	AAV57287 Human ORF2849
18	733	31.5	403	22	AAV78615 Human G-protein co
19	733	31.5	427	20	AAV32142 Human polypeptide,
20	732	31.5	403	22	AAV93548 Human RECAP poly
21	558	24.0	357	22	AAV33311 Peptide #2561 enco
22	558	24.0	357	22	AAB68891 Peptide #2594 enco
23	550.5	23.7	313	22	AAB29910 Protein #2505 enco
24	550.5	23.7	313	22	AAB35088 Human brain expres
25	550.5	23.7	313	22	AAB20506 Human bone marrow
26	550.5	23.7	313	22	AAV55909 Peptide #2536 enco
27	550.5	23.7	313	22	AAV68279 Peptide #2630 enco
28	550.5	23.7	313	22	AAV16102 Peptide #2511 enco
29	550.5	23.7	313	22	AAV03829 Rat skin cell tran
30	550.5	23.7	313	22	AAV76071 Skin cell protein,
31	533	22.9	347	21	AAB56010 Human orphan GPCR
32	533	22.9	347	22	AAV6010 Human G-protein co
33	524	22.5	317	22	AAB12274 Human G-protein co
34	524	22.5	362	22	AAB06764 Human secreted pro
35	518	22.3	323	22	AAV69174 Human ovarian tumo
36	478	20.6	106	20	AAV36337 Extended human sec
37	347	14.9	200	20	AAV76571 Human secreted pro
38	321	13.8	150	20	AAV36003 Human secreted pro
39	320	13.8	60	21	AAV00599 Human 5' EST secre
40	320	13.8	61	20	AAV11898 Hydrophobic domain
41	236	10.1	201	21	AAV78809 Human 5' EST secre
42	218	9.4	88	20	AAV11542 Human GPCR homolog
43	213	9.2	137	22	AAB12187 Human 5' EST secre
44	198	8.5	79	20	AAV11602 Human excretory re
45	188	8.1	119	22	AAV99779

ALIGNMENTS

RESULT 1
AAV57283
ID AAV57283 standard; Protein; 441 AA.
XX
AC AAV57283;
XX
DT 05-JUN-2000 (first entry)
XX
XX Human GPCR protein (HGPRP) sequence (clone ID 1258981).
DE
DE Human; G protein coupled protein receptor; HGPRP; cell proliferation;
KW neurological; immune disorder; cytostatic; anti-arteriosclerotic;
KW anti-atherosclerotic; hepatotropic; antiinflammatory; virucide; leukemia;
KW immunomodulatory; anemia; asthma; gastrointestinal; anti-epileptic;
KW anti-Alzheimer's; anti-Parkinsonian; gene therapy.
XX
XX Homo sapiens.
XX
PN WO200015793-A2.
XX
XX 23-MAR-2000.
XX
XX 17-SEP-1999; 99WO-US20958.
XX
XX 17-SEP-1998; 98US-0156513.
XX
XX (INCY-) INCYTE PHARM INC.
PI Bandman O, Lal P, Tang YT, Corley NC, Guegler KJ, Gorgone GA;
PI Baughn MR;
XX
XX WPI; 2000-271432/23.
DR N-PSDB; AA290521.
XX

PT Human G protein coupled protein receptor peptides useful for the
 PT prevention, diagnosis and treatment of cell proliferative, neurological
 PT and immune disorders -

PS Claim 1; Page 59-60; 71pp; English.

XX The invention provides human G protein coupled protein receptor (HGRP)
 CC polypeptides and polynucleotides encoding them. The polypeptides can be
 CC produced by standard recombinant methodology. The polynucleotides and
 CC polypeptides may be used in the prevention, treatment and diagnosis of
 CC diseases associated with their inappropriate expression. Diseases that
 CC can be treated are cell proliferative disorders (e.g. arteriosclerosis,
 CC atherosclerosis and hepatitis), cancers (e.g. leukemia, melanomas and
 CC adenocarcinoma), immune disorders (e.g. anemia, asthma and Crohn's
 CC disease) and neurological disorders (e.g. epilepsy, Alzheimer's disease
 CC and Parkinson's disease). The anti-HGRP antibodies may also be used as
 CC diagnostic agents for detecting the presence of HGRP polypeptides in
 CC samples (e.g. by enzyme linked immunosorbent assay (ELISA)). Sequences
 CC AAY57283-288 represent the HGRP polypeptides.

XX Sequence 441 AA;

Query Match 100.0%; Score 2326; DB 21; Length 441;
 Best Local Similarity 100.0%; Pred. No. 6.5e-241;
 Matches 441; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAIHKALVMCLGLPLFLPPGAWAGHVPVPGSCQGLNPLLYNLCDSGAWGIVLEAVAGAG 60
 Db 1 maihkalvmclglplflfpgawagghvppgcsqglnplynlnclsdsgawgivleavagag 60
 QY 61 IVTTFVLTIILVASLPFVQDTKRSLLGTVQVFFLLGLTGLFCLVFACVVKPDFSTCASRR 120
 Db 61 ivttfvtlilvaslpfvgdtkrsligtqvffllgtlglfclvfvacvvpkpdfstcasrr 120
 QY 121 FLFGVLFALFCSCLAHVFALNFKRNHGRGWVIFTVALLLTLVEVIINTEWLIITLV 180
 Db 121 flfgvlfalfcsciaahvfaInflarknhgrgwwiftvallltlveviintewliitlv 180
 QY 181 RGSCEGPGQNSAGWAVASPCAIANMDFVMALYVMLLLGLGAFLGAWPALCGRYKRWK 240
 Db 181 rgsceggpgqnsagwavaspcalnmdfvmallyvmllllglaflgawpalcgrykrrwk 240
 QY 241 HGVEVLLTTATSVAIWVIMYTYGNKQNSPTWDDPTLAIALAANAWAFVLYVPIEV 300
 Db 241 hgvevllttatsvaiwvImyTYGNKqnsptwddptlaialaanaWafvlyvIpev 300
 QY 301 SQVTKSPQSQYQGDMPYTRGVGYETILKEQKGSMFVENKAFSMDPEVAAKRPVSPYSG 360
 Db 301 sqvtkspqsgyqgdmpytrgvgyetilikEqkgsmfvenkafsmdepvAAKrpvsySG 360
 QY 361 YNGQLLTSVYOPTMALMHKVPSEGAVDIILPRATANSQVMSNSTLRAEDMYSAQSHQ 420
 Db 361 yngqltstvypTmalmhkvpsEGavdiilpratanSQVmsnstlraedmySaShq 420
 QY 421 AATPPKDGKNSQVFRNPVVD 441
 Db 421 aatppkdGknsqvfrnpvvd 441

RESULT 2

AAV50932
 ID AAV50932 standard; Protein; 441 AA.

XX AAY50932;

XX 10-MAR-2000 (first entry)

XX Human fetal brain cDNA clone vc26_1 derived protein #1.

XX Human; secreted protein; treatment; nutritional activity; cytokine;
 KW cell proliferation; cell differentiation; hematopoiesis regulation;
 KW tissue growth; activin; inhibin; chemotactic; chemokinetic; hemostatic;

thrombolytic; anti-inflammatory; invasion suppressor; tumor inhibition;
 gene therapy.

OS Homo sapiens.

PN WO995721-A1.

XX 04-NOV-1999.

XX 23-APR-1999; 99WO-US08504.

XX 24-APR-1998; 98US-0082904.

PR 11-JUN-1998; 98US-0088994.

PR 12-JUN-1998; 98US-0089278.

PR 02-JUL-1998; 98US-0091647.

PR 24-AUG-1998; 98US-0097639.

PR 22-APR-1999; 99US-0097639.

XX (ALPH-) ALPHAGENE INC.

XX Valenzuela D, Yuan O, Hoffman H, Hall J, Rapiejko P;

DR WPI: 2000-052801/04.

DR N-PSDB; AA243798.

XX New polynucleotides encoding secreted human proteins, derived from

PT human fetal brain, adult skin, adult brain, adult heart, adult thymus

XX and adult aorta cDNA libraries.

PS Claim 53a; Page 246-247; 282pp; English.

XX This invention describes novel human secreted proteins which are encoded
 CC by polynucleotides obtained from fetal brain, adult skin, adult brain,
 CC adult heart, adult thymus and adult aorta cDNA libraries. The
 CC polynucleotides and proteins are predicted to have biological activities
 CC which would make them suitable for treating, preventing or ameliorating
 CC medical conditions in humans and animals, although no supporting data
 CC is given. Suggested activities include nutritional activity, cytokine
 CC and cell proliferation/differentiation activity, immune stimulating
 CC (e.g. as vaccines) or suppressing activity, hematopoiesis regulating
 CC activity, tissue growth activity, activin/inhibin activity,
 CC chemotactic/chemokinetic activity, hemostatic and thrombolytic activity,
 CC receptor/ligand activity, anti-inflammatory activity, cadherin/tumor
 CC invasion suppressor activity, and tumor inhibition activity. The
 CC polynucleotides are also stated to be useful for gene therapy.

CC AAY50905-Y50947 represent the secreted proteins described in the method
 CC of the invention which are encoded by the polynucleotides represented in
 CC AA243777-243808.

XX Sequence 441 AA;

Query Match 100.0%; Score 2326; DB 21; Length 441;
 Best Local Similarity 100.0%; Pred. No. 6.5e-241;
 Matches 441; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAIHKALVMCLGLPLFLPPGAWAGHVPVPGSCQGLNPLLYNLCDSGAWGIVLEAVAGAG 60
 Db 1 maihkalvmclglplflfpgawagghvppgcsqglnplynlnclsdsgawgivleavagag 60
 QY 61 IVTTFVLTIILVASLPFVQDTKRSLLGTVQVFFLLGLTGLFCLVFACVVKPDFSTCASRR 120
 Db 61 ivttfvtlilvaslpfvgdtkrsligtqvffllgtlglfclvfvacvvpkpdfstcasrr 120
 QY 121 FLFGVLFALFCSCLAHVFALNFKRNHGRGWVIFTVALLLTLVEVIINTEWLIITLV 180
 Db 121 flfgvlfalfcsciaahvfaInflarknhgrgwwiftvallltlveviintewliitlv 180
 QY 181 RGSCEGPGQNSAGWAVASPCAIANMDFVMALYVMLLLGLGAFLGAWPALCGRYKRWK 240
 Db 181 rgsceggpgqnsagwavaspcalnmdfvmallyvmllllglaflgawpalcgrykrrwk 240
 QY 241 HGVEVLLTTATSVAIWVIMYTYGNKQNSPTWDDPTLAIALAANAWAFVLYVPIEV 300

|||||
241 hgfvllttatstvalwvmytygnkqnsptdptlaialaanaawafvfyvipev 300
QY 301 SQVTKSSPEQSYQGDYMPTRGVGYETILKEQKQSMFVENKAFSMDPEPVAARPVSPYSQ 360
Db 301 sqvtksspeqsyqgdmyptrgvgyetilkqkqsgmfvenkafsmdepvaakrpvpsysq 360
QY 361 YNGQLTTSVYQPTTEMALMHKVPSEGAYDIILPRATANSQVMSGANSTLRAEDMYSAQSHQ 420
Db 361 yngqlttsvyqptemalmhkvpssegaydiilpratansqvmgsanstlraedmysaqshq 420
QY 421 AATPPKDGKNSQVFRNPYVWD 441
Db 421 aatppkdgknsqvrnpyvwd 441
RESULT 3
AAM39849
ID AAM39849 standard; Protein; 441 AA.
XX
AC AAM39849;
XX
XX
DT 22-OCT-2001 (first entry)
XX
XX Human polypeptide SEQ ID NO 2994.
DE
DE Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
XX Homo sapiens.
OS
XX WO200153312-A1.
PN
XX
XX
PD 26-JUL-2001.
XX
XX 26-DEC-2000; 2000WO-US34263.
XX
XX 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI: 2001-442253/47.
DR N-PSDB; AA159005.
DR
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
PT
XX
XX Example 4; SEQ ID NO 2994; 10078pp; English.
PS
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM38642-AA42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the

CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemia and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 441 AA;
Query Match 100.0%; Score 2326; DB 22; Length 441;
Best Local Similarity 100.0%; Pred. No. 6.5e-241;
Matches 441; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATHKALVMCLGLPLFLPGAWAQGHVPPCCSQGLNPLYYNLCDRSGANGVLEAVAGAG 60
Db 1 maithkalvmclglplflpgawaqghvppccsglnplynlncdrsgawgvlavagag 60
QY 61 IVTTFVLTITILVASLPFVQDTKKRSLLTGTQVFFLLGTGLFCLVFACVKVPDFSTCASRR 120
Db 61 ivttfvtltilvaslpfvqdtkkrsllgtqvfllgtglfclvfvacvvpkdfstcasrr 120
QY 121 FLFGVLFAICFSCLAARHVFALNFLARKNHGPRGWIFTVALLTLTVEVIINTEWLIITLV 180
Db 121 flfgvlfaicfscslaahvfalnflarknhgprgwiiftvalltlveviintewliitlv 180
QY 181 RGSCEGPGQNSAGWAVASPCAIAANMDFVMALIYYVMLLLGAFLGAWPALCGRYKRWK 240
Db 181 rgsgegpgqngnsagwavaspcaiamdmfvmaliyvmllllgaflgawpalcgryrkrwk 240
QY 241 HGVEVLLTTSVVAIVWVIVMYTYGNKOHNSPTWDDPTLAIALANAWAFVIFYVPEV 300
Db 241 hgvevllttatsvalwvmytygnkqnsptwddptlaialaanaawafvifyvipev 300
QY 301 SQVTKSSPEQSYQGDYMPTRGVGYETILKEQKQSMFVENKAFSMDPEPVAARPVSPYSQ 360
Db 301 sqvtksspeqsyqgdmyptrgvgyetilkqkqsgmfvenkafsmdepvaakrpvpsysq 360
QY 361 YNGQLTTSVYQPTTEMALMHKVPSEGAYDIILPRATANSQVMSGANSTLRAEDMYSAQSHQ 420
Db 361 yngqlttsvyqptemalmhkvpssegaydiilpratansqvmgsanstlraedmysaqshq 420
QY 421 AATPPKDGKNSQVFRNPYVWD 441
Db 421 aatppkdgknsqvrnpyvwd 441
RESULT 4
AAB88591
ID AAB88591 standard; Protein; 441 AA.
XX
XX AAB88591;
XX
XX 04-JUN-2001 (first entry)
XX
XX Human hydrophobic domain containing protein clone HP10704 #95.
XX
XX Human; hydrophobic domain; immunosuppressant; anti-HIV; neuroprotective;
KW antianemic; vulnery; antiulcer; osteopathic; anti-inflammatory;
KW cytostatic; gene therapy; autoimmune disorder; multiple sclerosis;
KW HIV infection; anaemia; burn; ulcer; osteoporosis; tumour; wound healing;
KW inflammatory bowel disease; nutritional supplement; appetite; vaccine;
KW behavioural characteristic; immune response.
XX
XX Homo sapiens.
OS
XX
XX WO200112660-A2.
XX
XX 22-FEB-2001.
PD
XX
XX 10-AUG-2000; 2000WO-JP05356.
PF
XX

CC arthritis and diabetes.

XX Sequence 441 AA;

Query Match 100.0%; Score 2326; DB 22; Length 441;
Best Local Similarity 100.0%; Pred. No. 6.5e-241;
Matches 441; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAIHKALVMCLGLPLFLPGAWAQGHVPPGCSQGLNPLYNLCDSRSGAWGIVLEAVAG 60
Db 1 maihkalmcglplflfpgawagqhvpvpcsgqlnplynldrsdgawgivleavag 60
Qy 61 IVTTFVLTIIIVASLPFVQDTKKRSLTGTQVFFLLGLFCLVFAVCWKPDPFSTCASRR 120
Db 61 ivttfvltiilvaslpfvqdtkkrrslgtqvffllgtlglfclvfcvckpdpfstcasrr 120
Qy 121 FLFGVLFAICFSCIAAHVFALNFILARKNHGRGWIFTVALLTLVEIINTEWLIITLV 180
Db 121 flfgvlfaicfsciaahvfalnflarknhgrgwwiftvalltlciveiintewliitlv 180
Qy 181 RGSCEGPGQNSSAGWAVASPCAIANMDFVMAIYVMLLLGLGAWPALCGRYKRWK 240
Db 181 rgsceggpgqnsagwavaspcailanmdfvmaliyvmlllglgawpalcgryrkwrk 240
Qy 241 HGVFVLLTATSVAIWVWVIMVYTYGNKHNSPTWDDPTLAIALANAWAFVIFYPIEV 300
Db 241 hgvflltatsvaivwvviwmytygnkhnspwddptlaialaanawafvifypiev 300
Qy 301 SQVTKSSPEQSYQGDMPYTRGVGYETILKEQKQSMFVENKAFSDPEVAAKRPVPSYG 360
Db 301 sqvtksspeqsygqdmyptrgvgyetilkckqgsmfvenkafsdmdepvaakrpvpsyg 360
Qy 361 YNGQLLTSVYQPTMALMHKVPSEGAYDIILPRATANSQVNGSANSILRAEDMYSAQSHQ 420
Db 361 yngqlttsvyqptemalmhkvpssegaydiilpratansqvmgsansilraedmysaqshq 420
Qy 421 AATPPKDGKNSQVFRNPYVWD 441
Db 421 aatppkdgknsqvfrnpyvwd 441

RESULT 6

AAU14166
ID AAU14166 standard; Protein; 486 AA.

AC AAU14166;

XX 24-OCT-2001 (first entry)

XX Human novel protein #37.

XX Human: novel protein; Antianemic; osteopathic; antiinflammatory;
KW immunomodulatory; cytostatic; neuroprotective; vulnerary; nootropic;
KW anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral;
KW antibacterial; antiallergic; dermatological; haemostatic; antiasthmatic;
KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;
KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;
KW tissue regeneration; immune disorder.

XX Homo sapiens.

XX OS

XX PN

XX WO200155437-A2.

XX XX

PD 02-AUG-2001.

XX 25-JAN-2001; 2001WO-US02623.

XX 25-JAN-2000; 2000US-0491404.

XX (HYSE-) HYSEQ INC.

XX PA

XX Tang YT, Liu C, Drmanac RT;

XX PI

XX

DR

DR

XX

PT WPI: 2001-451939/48.

PT N-PSDB; AAS22471.

XX

PS Isolated polypeptides useful for treating anti-inflammatory diseases,

PS nervous system disorders, and for regenerating bone and cartilage -

XX

XX Example 4; Page 545-546; 894pp; English.

XX

CC The invention relates to polynucleotides encoding novel human
CC proteins or their active domains. The polypeptides, polynucleotides and
CC antibodies raised against the polypeptides are used in a method of
CC treatment of a mammal and prevention of disorders caused by the aberrant
CC protein expression or activity. The polypeptides can be used as
CC molecular weight markers, food supplements, and in antibody production.
CC The polypeptides are used to identify compounds which bind to the
CC polypeptides. Polynucleotides of the invention are used as probes and
CC primers, for sequencing, for chromosome or gene mapping, in the
CC production of recombinant proteins, and in generating anti-sense DNA or
CC RNA and in gene therapy. Polypeptides of the invention can be used to
CC target drugs to a tumour, in assays to determine biological activity, to
CC raise antibodies/ elicit an immune response, to determine quantitative
CC protein levels, as tissue markers, and to isolate receptors or ligands.
CC Polypeptides of the invention may also be useful in treating platelet
CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,
CC ligament and/or nerve tissue, wound healing, treating burns, promoting
CC the proliferation, differentiation and survival of stem cells, as a
CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,
CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral
CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
CC fungal infection or from autoimmunity, cancer, allergy, asthma,
CC graft-versus-host disease, eczema, haemophilia, thrombosis,
CC anti-inflammatory diseases, nervous system disorders, and infection.
CC The present sequence represents a protein of the invention.

XX Sequence 486 AA;

Query Match 100.0%; Score 2326; DB 22; Length 486;

Best Local Similarity 100.0%; Pred. No. 7.6e-241;

Matches 441; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAIHKALVMCLGLPLFLPGAWAQGHVPPGCSQGLNPLYNLCDSRSGAWGIVLEAVAG 60

Db 46 maihkalmcglplflfpgawagqhvpvpcsgqlnplynldrsdgawgivleavag 105

Qy 61 IVTTFVLTIIIVASLPFVQDTKKRSLTGTQVFFLLGLFCLVFAVCWKPDPFSTCASRR 120

Db 106 ivttfvltiilvaslpfvqdtkkrrslgtqvffllgtlglfclvfcvckpdpfstcasrr 165

Qy 121 FLFGVLFAICFSCIAAHVFALNFILARKNHGRGWIFTVALLTLVEIINTEWLIITLV 180

Db 166 flfgvlfaicfsciaahvfalnflarknhgrgwwiftvalltlciveiintewliitlv 225

Qy 181 RGSCEGPGQNSSAGWAVASPCAIANMDFVMAIYVMLLLGLGAWPALCGRYKRWK 240

Db 226 rgsceggpgqnsagwavaspcailanmdfvmaliyvmlllglgawpalcgryrkwrk 285

Qy 241 HGVFVLLTATSVAIWVWVIMVYTYGNKHNSPTWDDPTLAIALANAWAFVIFYPIEV 300

Db 286 hgvflltatsvaivwvviwmytygnkhnspwddptlaialaanawafvifypiev 345

Qy 301 SQVTKSSPEQSYQGDMPYTRGVGYETILKEQKQSMFVENKAFSDPEVAAKRPVPSYG 360

Db 346 sqvtksspeqsygqdmyptrgvgyetilkckqgsmfvenkafsdmdepvaakrpvpsyg 405

Qy 361 YNGQLLTSVYQPTMALMHKVPSEGAYDIILPRATANSQVNGSANSILRAEDMYSAQSHQ 420

Db 406 yngqlttsvyqptemalmhkvpssegaydiilpratansqvmgsansilraedmysaqshq 465

Qy 421 AATPPKDGKNSQVFRNPYVWD 441

Db 466 aatppkdgknsqvfrnpyvwd 486

```
RESULT 7
AA49156
ID AA49156 standard; Protein; 441 AA.
XX
AC AA49156;
XX
DT 17-JAN-2000 (first entry)
XX
DE G-protein coupled receptor AXOR4 amino acid sequence #1.
XX
KW AXOR4; G-protein coupled receptor; RAIG1; pain; cancer; diabetes;
KW obesity; anorexia; bulimia; asthma; stroke; ulcer; allergy; infertility;
KW neurological disorder; schizophrenia.
XX
OS Homo sapiens.
XX
PN W09953054-A1.
XX
PD 21-OCT-1999.
XX
PF 07-APR-1999; 99WO-GB01067.
XX
PR 08-APR-1998; 98GB-0007723.
XX
PA (SMIK ) SMITHKLINE BEECHAM PLC.
XX
PI Medhurst A, Michalovich D, Pangalos M, Hill J;
XX
DR WPI; 1999-620425/53.
XX
DR N-PSDB; AAZ32501.
XX
PT New AXOR4 G-protein coupled receptor polypeptides and related nucleic
PT acid, for treating, preventing or diagnosing e.g. cancer
XX
PS Claim 2; Page 38-39; Sipp; English.
XX
CC This is the G-protein coupled receptor AXOR4, amino acid sequence #1. The
CC AXOR4 polypeptide has homology with human putative G-protein coupled
CC receptor RAIG1. The AXOR4 polypeptide and polynucleotide sequences can be
CC used in the production of antibodies specific for AXOR4. The sequences
CC may also be used to screen for AXOR4 agonists or antagonists.
CC Alternatively the effect of a candidate agonist or antagonist compound on
CC the production of mRNA encoding AXOR4 may be detected using an ELISA
CC assay. Diseases or conditions arising from altered expression or activity
CC of AXOR4 may be diagnosed by detecting the AXOR4 protein in a sample from
CC a patient or detecting a mutation in the AXOR4 nucleotide sequence in the
CC genome of a patient. These diseases or conditions include pain, cancer,
CC diabetes, obesity, anorexia, bulimia, asthma, hypo- or hypertension,
CC stroke, ulcers, allergy, benign prostatic hypertrophy, migraine,
CC epilepsy, vomiting, psychosis, infertility, and neurological disorders
CC (e.g. anxiety, schizophrenia, depression, delirium, dementia,
CC neurodegeneration or severe mental retardation). Detection of the AXOR4
CC nucleotide sequence may be used for chromosome mapping and tissue
CC localization. The polynucleotide and polypeptide sequences may be
CC administered to patients as vaccines or as part of a gene therapy regime
CC respectively, to treat the diseases and conditions listed above.
XX
SQ Sequence 441 AA;
```

```
Query Match 100.0%; Score 2325; DB 20; Length 441;
Best Local Similarity 99.88; Pred. No. 8.4e-241;
Matches 440; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAIHKALVCMGLPLFLFPGWAQHVPVPGSCQGLNPLYNLCDRSGAWGIVLEAVAGAG 60
Dy 1 maihkalvcmglplflfpgwaaghvppgcsqglnplynlncldrsgawgivleavag 60
Qy 61 IVTTFVLTIIIVASLPFVQDTKKRSLGTQVFFLLGTGLGLFCLVACVVKPDFSTCSARR 120
Dy 61 ivttfvltiilvaslptfvqdtkkrrslgtqvffllgtlglfclvfacvvpkpdfstcsarr 120
```

```
Qy 121 FLFGVLFALCFSCLAHVFALNFLARKNHGRGWIFVTALLTLVEIINTEWLIITLV 180
Dy 121 flfgvlfalcfsclaahvfalnflarknhgrgwwifvtalltlveintewliitlv 180
Qy 181 RGSSEGGPQGNSSAGWAVASPCAIANMDFVMALYVVMILLGALGAWPALCGRYKRWK 240
Dy 181 rgsgseggpqgnssagwavaspcaianmdfvmaliyvvmillgafgawpalcgrkwrk 240
Qy 241 HGVFVLLTATSAIVWVIMVYTYGNKHNSPTWDDPTLAIALAANAWAFVLYVIVPEV 300
Dy 241 hgvfvlltatsaivwvwmvtytygnkhnsptwddptlaialaanaawafvlyvipev 300
Qy 301 SOVTKSSPEQSYQGDMPYTRGVYETILKEQKQSMFVKNFAFMDPEVAAKRPVSPYSG 360
Dy 301 sovtksspeqsyqgdmpytrgvgyetilkeqkqsmfvenkafmdpevvaakrpvspysg 360
Qy 361 YNGQLTTSVYQPTMALMHKVPSEGYDIILPRATANSQVMSANSTLRAEDMYSQASHQ 420
Dy 361 yngqlttsvyqptemalmhkvpssegaydiilpratansqvmgsanstlraedmysaqshq 420
Qy 421 AATPPKDGKNSQVFRNPYVWD 441
Dy 421 aatppkdgknsqvfrnpyvwd 441

RESULT 8
AAM41635
ID AAM41635 standard; Protein; 460 AA.
XX
AC AAM41635;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 6566.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
PN W0200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
XX
PR 25-APR-2000; 2000US-0523317.
XX
PR 09-JUL-2000; 2000US-0598042.
XX
PR 19-JUL-2000; 2000US-0620312.
XX
PR 03-AUG-2000; 2000US-0653450.
XX
PR 14-SEP-2000; 2000US-0662191.
XX
PR 19-OCT-2000; 2000US-0693036.
XX
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
XX
DR N-PSDB; AAI60791.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries
XX
PS Example 2; SEQ ID NO 6566; 10078pp; English.
```


XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAW42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 460 AA;

Query Match 99.7%; Score 2320; DB 22; Length 460;
Best Local Similarity 99.8%; Pred. No. 3.1e-240;
Matches 440; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAIHKALVCMCLGLPLFLFPAGAWAGHVPVPGCSQGLNPLYNLCDSRGAWGIVLEAVAGAG 60
Db 20 MAIHKALVCMCLGLPLFLFPAGAWAGHVPVPGCSQGLNPLYNLCDSRGAWGIVLEAVAGAG 79

QY 61 IVTTFVLTIIIVASLPFVQDTKKRSLGTQVFFLLGLTGLFCLVFCACVVPDFSTCASRR 120
Db 80 IVTTFVLTIIIVASLPFVQDTKKRSLGTQVFFLLGLTGLFCLVFCACVVPDFSTCASRR 139

QY 121 FLFGVLFAICFSCLAHVFALNFALRNKHGPRGWVFTVALLTLVEVINTEWLIITLV 180
Db 140 FLFGVLFAICFSCLAHVFALNFALRNKHGPRGWVFTVALLTLVEVINTEWLIITLV 199

QY 181 RGSSEGGPQGNSSAGWAVASPCAIANMDFVMALIVYMLLLGAFLGAWPALCGRYKRWK 240
Db 200 RGSSEGGPQGNSSAGWAVASPCAIANMDFVMALIVYMLLLGAFLGAWPALCGRYKRWK 259

QY 241 HGVEVLLTATSVAIWWVIMTYGNKQHNPSPTDDPTLAIALAANAFAVLPVPIPEV 300
Db 260 HGVEVLLTATSVAIWWVIMTYGNKQHNPSPTDDPTLAIALAANAFAVLPVPIPEV 319

QY 301 SQVTKSSPEQSYQDMYPTRCGVETILKEQKGSMFVENKAFSMDPEPVAARKRVPSPYSG 360
Db 320 SQVTKSSPEQSYQDMYPTRCGVETILKEQKGSMFVENKAFSMDPEPVAARKRVPSPYSG 379

QY 361 YNGOLLTSVYQPTMALMHKVPSEGYDIIILPRATANSQVNGSANSSTLRADMYSAQSHQ 420
Db 380 YNGOLLTSVYQPTMALMHKVPSEGYDIIILPRATANSQVNGSANSSTLRADMYSAQSHQ 439

QY 421 AATPPKDGKNSQVFRNPYVMD 441
Db 440 AATPPKDGKNSQVFRNPYVMD 460

RESULT 9
ID AAY44273 standard; Protein; 479 AA.
XX
AC AAY44273;
XX
DT 28-FEB-2000 (first entry)
XX
DE Human Metabotropic Glutamate Receptor-like protein, MGRcm.
XX
KW Metabotropic Glutamate Receptor-like protein; MGLuR; MGRcm; G-protein;
KW 7-transmembrane; GABA receptor; diagnostic marker; pharmacological agent;
KW central and peripheral nervous system; cardiac disorder; treatment;
KW urologic disorder; gastrointestinal disorder; diagnosis; seizure; anoxia;
KW epilepsy; Alzheimer's disease; Huntington's disease; Parkinson's disease;

KW antibody; inhibitor; screening.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 20..40 /label= Extracellular_domain
FT /note= "Relatively short"
FT Domain 40..300 /note= "Group of 7 transmembrane segments"
FT Domain 301..479 /label= C-terminal_cytosolic_domain
FT /note= "Relatively long"
PN WO9960121-A1.
XX
PD 25-NOV-1999.
XX
XX 19-MAY-1999; 99WO-IL00265.
XX
XX 19-MAY-1998; 98US-0085973.
XX (COMP-) COMPUGEN LTD.
XX Mintz L, Savitsky K, Toporik A;
PI
XX WPI; 2000-086596/07.
DR N-PSDB; AA29292.
XX
PT New metabotropic glutamate receptor-like protein, useful for, e.g.
PT diagnosing neurological diseases
XX
PS Claim 1; Fig 1B; 84pp; English.
XX
CC The present amino acid sequence is the human metabotropic
CC glutamate receptor like protein (MGLuR-like receptor protein), MGRcm. It
CC is a novel member of the metabotropic glutamate/GABA receptor class of
CC 7-transmembrane G-protein coupled receptors showing sequence similarity.
CC It is responsible for activation of intracellular G-proteins. The MGRcm
CC receptor protein and encoding nucleic acid sequence, are useful as
CC diagnostic markers and pharmaceutical agents for various diseases of the
CC central and peripheral nervous system, as well as cardiac, urologic and
CC gastrointestinal disorders. It can be used for the diagnosis and
CC treatment of neurological and neurodegenerative disorders, such as
CC seizures, epilepsy, anoxia, Alzheimer's disease, Huntington's disease,
CC Parkinson's disease, etc. The antibody specific against MGRcm has
CC diagnostic and therapeutic applications against these disease states.
CC This receptor protein can also be employed for screening of potential
CC peptide or small molecule inhibitors of the relevant receptor-ligand
CC interaction.
XX
SQ Sequence 479 AA;

Query Match 95.7%; Score 2226; DB 21; Length 479;
Best Local Similarity 100.0%; Pred. No. 4.2e-230;
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAIHKALVCMCLGLPLFLFPAGAWAGHVPVPGCSQGLNPLYNLCDSRGAWGIVLEAVAGAG 60
Db 46 MAIHKALVCMCLGLPLFLFPAGAWAGHVPVPGCSQGLNPLYNLCDSRGAWGIVLEAVAGAG 105

QY 61 IVTTFVLTIIIVASLPFVQDTKKRSLGTQVFFLLGLTGLFCLVFCACVVPDFSTCASRR 120
Db 106 IVTTFVLTIIIVASLPFVQDTKKRSLGTQVFFLLGLTGLFCLVFCACVVPDFSTCASRR 165

QY 121 FLFGVLFAICFSCLAHVFALNFALRNKHGPRGWVFTVALLTLVEVINTEWLIITLV 180
Db 166 FLFGVLFAICFSCLAHVFALNFALRNKHGPRGWVFTVALLTLVEVINTEWLIITLV 225

QY 181 RGSSEGGPQGNSSAGWAVASPCAIANMDFVMALIVYMLLLGAFLGAWPALCGRYKRWK 240
Db 226 RGSSEGGPQGNSSAGWAVASPCAIANMDFVMALIVYMLLLGAFLGAWPALCGRYKRWK 285

[illegible]

CC the proliferation, differentiation and survival of stem cells, as a
 CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,
 CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral
 CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
 CC fungal infection or from autoimmunity, cancer, allergy, asthma,
 CC graft-versus-host disease, eczema, haemophilia, thrombosis,
 CC anti-inflammatory diseases, nervous system disorders, and infection.
 CC The present sequence represents a protein of the invention.

XX Sequence 231 AA;

Query Match 52.5%; Score 1221; DB 22; Length 231;
 Best Local Similarity 100.0%; Pred. No. 1.4e-122;
 Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 211 MALIYVMLLLGAFGLGAWPALCGRYKRWKRGVFLVLTATTSVAIVWVWVWYTGKQH 270
 Db 1 maliyvmlllgafglgawpalcgrkwrkhgfvllttatsvalwvwmvmytygnkqh 60
 QY 271 NSPTWDDPTLAIALAANAWAFVLPVPIEVSQVTKSSPEOSYQGDMPYTRGVGYETILKE 330
 Db 61 nsptwdptlaialaanaawafvlpvpievsqvtckspedsgyqgdmyptrgvgyetilke 120
 QY 331 QKGSMFVENKAFSMDEPVAARKRPVSPSYNGQLTSVYQPTMALMHKVPSEGYDII 390
 Db 121 qkgsmfvenkafsmdepvaarkrpvspysyngqltsvyqptemalmhkvpssegaydii 180
 QY 391 LPRATANSQVNGSANSFLRADMYSAQSHQAATPPKDGKNSQVFRNPYVWD 441
 Db 181 lpratansqvmgsanstiraedmysaqshqaatppkdgknsqvrnpvwd 231

RESULT 13

ID AAY44274 standard; Protein; 296 AA.

AC AAY44274;

DT 28-FEB-2000 (first entry)

DE Mouse Metabotropic Glutamate Receptor-like protein, MGRcm.

XX Metabotropic Glutamate Receptor-like protein; MGLuR; MGRcm; G-protein;
 KW 7-transmembrane; GABA receptor; diagnostic marker; pharmaceutical agent;
 KW central and peripheral nervous system; cardiac disorder; treatment;
 KW urologic disorder; gastrointestinal disorder; diagnosis; seizure; anoxia;
 KW epilepsy; Alzheimer's disease; Huntington's disease; Parkinson's disease;
 KW antibody; inhibitor; screening; mouse; homologue.

XX Mus musculus.

PH Key Location/Qualifiers

FT Protein 1..296

FT /label= Mouse_MGRcm_protein

FT /note= "Homolog of human MGRcm"

FT Misc-difference 37 /note= "Corresponds to CYA codon"

FT Misc-difference 148 /note= "Corresponds to TWV codon"

FT Misc-difference 155 /note= "Corresponds to WGC codon"

FT /note= "Corresponds to WGC codon"

PN WO960121-A1.

XX 25-NOV-1999.

XX 19-MAY-1999; 99WO-IL00265.

XX 19-MAY-1998; 98US-0085973.

XX (COMP-) COMPUGEN LTD.

XX

PI Mintz L, Savitsky K, Toporik A;

XX WPI: 2000-086596/07.

DR N-PSDB; AAZ29293.

XX New metabotropic glutamate receptor-like protein, useful for, e.g.

PT diagnosing neurological diseases -

XX Claim 1; Fig 2B; 84pp; English.

XX The present amino acid sequence is the mouse metabotropic

CC glutamate receptor like protein (MGLuR-like receptor protein), MGRcm. It

CC is a novel member of the metabotropic glutamate/GABA receptor class of

CC 7-transmembrane G-protein coupled receptors showing sequence similarity.

CC It is homologous to the human MGRcm sequence.

CC It is responsible for activation of intracellular G-proteins. The MGRcm

CC receptor protein and encoding nucleic acid sequence, are useful as

CC diagnostic markers and pharmaceutical agents for various diseases of the

CC central and peripheral nervous system, as well as cardiac, urologic and

CC gastrointestinal disorders. It can be used for the diagnosis and

CC treatment of neurological and neurodegenerative disorders, such as

CC seizures, epilepsy, anoxia, Alzheimer's disease, Huntington's disease,

CC Parkinson's disease, etc. The antibody specific against MGRcm has

CC diagnostic and therapeutic applications against these disease states.

CC This receptor protein can also be employed for screening of potential

CC peptide or small molecule inhibitors of the relevant receptor-ligand

CC interaction.

XX Sequence 296 AA;

SQ

Query Match 51.6%; Score 1200.5; DB 21; Length 296;

Best Local Similarity 84.9%; Pred. No. 3.4e-120;

Matches 230; Conservative 11; Mismatches 29; Indels 1; Gaps 1;

QY 1 MATHKALVMCLGLPLFLFPGAWAQGHVPPGCSQGLNPLYYNLCDRSGAWGIVLEAVAG 60

Db 27 mathrtlmcxgplf-fpgalaqnahppgcspldplynldcrgawgivleavag 85

QY 61 IVTFVLTITILVASLPEVQDTKRSLGTVQVFFLLGLFLCLVFCVAVKPDFSTCASRR 120

Db 86 iitfvltitilvasipfvqdtkrslgtqvfllgtlglfcvavkpdfstcasrr 145

QY 121 FLFGVLFPAICFSLAAHVFALNFLARKNHGPRGVITVALLTLVEVIINTEWLIITLV 180

Db 146 flxgvlfaixfscvlavhslinfiltrknhprgwwitvavllclveviintewliitlv 205

QY 181 RGSGEQPGQNSSAGWAVASPCATANNDFVMAIYVMLLLGAFGLGAWPALCGRYKRWK 240

Db 206 pggsgvqspianvsadstmtspcalanmdfvmaliyvmlllltaflgawptlccrkrwrk 265

QY 241 HGVFVLLTTATTSVAIVWVWVWVWYTGKQHN 271

Db 266 hgvfvlittvtsiaivwvwmvmytygneqhn 296

RESULT 14

AAY50933

ID AAY50933 standard; Protein; 302 AA.

XX AAY50933;

XX 10-MAR-2000 (first entry)

XX Human fetal brain cDNA clone vc26_1 derived protein #2.

XX Human; secreted protein; treatment; nutritional activity; cytokine;

XX cell proliferation; cell differentiation; hematopoiesis regulation;

XX tissue growth; activin; inhibitor; chemotactic; chemokinetic; hemostatic;

XX thrombolytic; anti-inflammatory; invasion suppressor; tumor inhibition;

XX gene therapy.

XX Homo sapiens.

XX

XX WO9955721-A1.
 XX 04-NOV-1999.
 XX 23-APR-1999; 99WO-US08504.
 XX 24-APR-1998; 98US-0082904.
 XX 11-JUN-1998; 98US-0088994.
 XX 12-JUN-1998; 98US-0089278.
 XX 02-JUL-1998; 98US-0091647.
 XX 24-AUG-1998; 98US-0097639.
 XX 22-APR-1999; 99US-0097639.
 XX (ALPH-) ALPHAGENE INC.
 XX Valenzuela D, Yuan O, Hoffman H, Hall J, Rapiejko P;
 XX WPI; 2000-052801/04.
 XX N-PSDB; AA243798.
 XX New polynucleotides encoding secreted human proteins, derived from
 PT human fetal brain, adult skin, adult brain, adult heart, adult thymus
 PT and adult aorta cDNA libraries.
 XX
 XX Disclosure; Page 274-275; 282pp; English.
 XX
 XX This invention describes novel human secreted proteins which are encoded
 CC by polynucleotides obtained from fetal brain, adult skin, adult brain,
 CC adult heart, adult thymus and adult aorta cDNA libraries. The
 CC polynucleotides and proteins are predicted to have biological activities
 CC which would make them suitable for treating, preventing or ameliorating
 CC medical conditions in humans and animals, although no supporting data
 CC is given. Suggested activities include nutritional activity, cytokine
 CC and cell proliferation/differentiation activity, immune stimulating
 CC (e.g. as vaccines) or suppressing activity, hematopoiesis regulating
 CC activity, tissue growth activity, activin/inhibin activity,
 CC chemotactic/chemokinetic activity, hemostatic and thrombolytic activity,
 CC receptor/ligand activity, anti-inflammatory activity, cadherin/tumor
 CC invasion suppressor activity, and tumor inhibition activity. The
 CC polynucleotides are also stated to be useful for gene therapy.
 CC AA30905-Y30947 represent the secreted proteins described in the method
 CC of the invention which are encoded by the polynucleotides represented in
 CC AA243777-243808.
 XX
 XX Sequence 302 AA;
 CC
 CC Query Match 43.9%; Score 1020.5; DB 21; Length 302;
 CC Best Local Similarity 66.6%; Pred. No. 7.8e-101;
 CC Matches 209; Conservative 10; Mismatches 30; Indels 65; Gaps 5;
 QY 1 MATHKALVMCLGLPLFLFPGAWAQGHVPGCCSGLNPLYYNLCDSRGANGIVLEAVAG 60
 DB 1 mathkalvmclglplflfgawagghvpgccsglnplynlnldrsgangivleavag 60
 QY 61 IVTFVTLTILVASLPVQDTKRSLTGQVFFLLGLFGLVFACVKVPDFSTCASRR 120
 DB 61 ivtftvtilvaslpfvqdtkrslrgtqvfllglgflvfcvkvkpdfstcasrr 120
 QY 121 FLFGVLFAICFSCLAARFALNFALRNHGRGWGVIFFVALLTLVEVIINTEWLIITLV 180
 DB 121 flfgvlfaicfscslaahvafalnflarknhgrgwgvi fvaalltlveviintewliitlv 180
 QY 181 RGSGBGPQGNSSA-----GWAVASPCAIANMDFVM 211
 DB 181 rsgsgbgpqpqgnssaaqgppvpptwlswhstscvccwpswpgpbcvvaats---- 236
 QY 212 ALIYVMLLLLGALFGAWPALCGRYK-----RWRKHGVFVLLTTATSVAILWVIVMTY 265
 DB 237 -----agvsmg---sicssqpqpplpygwcgssclitatssttvpqgm----- 276
 QY 266 GNQKHSPTWDDPT 279

DB 277 -----tpwpsps 284
 RESULT 15
 AA32141
 ID AAY32141 standard; Protein; 403 AA.
 XX
 AC AAY32141;
 XX
 DT 01-FEB-2000 (first entry)
 XX
 DE Human G-protein coupled receptor AXOR3.
 XX
 KW AXOR3; G-protein coupled receptor; human; pain; cancer;
 KW diabetes; obesity; anorexia; bulimia; asthma; hypotension;
 KW hypertension; stroke; ulcer; allergy; benign prostatic hypertrophy;
 KW migraine; epilepsy; vomiting; psychosis; infertility;
 KW neurological disorder; anxiety; schizophrenia; manic depression;
 KW depression; delirium; dementia; neurodegenerative disease;
 KW mental retardation; therapy; diagnosis.
 XX
 OS Homo sapiens.
 XX
 PN WO9952944-A1.
 XX
 PD 21-OCT-1999.
 XX
 PF 06-APR-1999; 99WO-GB01038.
 XX
 PR 08-APR-1998; 98GB-0007722.
 XX
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 XX
 XX Medhurst A, Michalovich D, Pangalos M, Hill J;
 DR WPI; 1999-633823/54.
 DR N-PSDB; AA220297.
 XX
 PT New isolated G-protein coupled receptor AXOR3 polypeptides, used to
 PT develop products for treating, e.g. pain, cancers, neurological and
 PT neurodegenerative disorders.
 XX
 PS Claim 1; Page 40-41; 51pp; English.
 XX
 CC The present sequence represents AXOR3, a novel human G-protein
 CC coupled receptor that shows homology to the human G-protein coupled
 CC receptor RAIG1. The invention relates to AXOR3 polypeptides and
 CC polynucleotides, recombinant materials and methods for their
 CC production. It also relates to methods for using such polypeptides
 CC and polynucleotides for treatment of pain, cancers, diabetes,
 CC obesity, anorexia, bulimia, asthma, hypotension, hypertension,
 CC stroke, ulcers, allergies, benign prostatic hypertrophy, migraine,
 CC epilepsy, vomiting, psychosis, infertility, neurological disorders
 CC including anxiety, schizophrenia, manic depression, depression,
 CC delirium, dementia, neurodegenerative diseases and severe mental
 CC retardation. Also disclosed are methods for identifying agonists
 CC and antagonists/inhibitors, and treating conditions associated with
 CC AXOR3 imbalance with the identified compounds. The invention also
 CC relates to diagnostic assays for detecting diseases associated with
 CC inappropriate AXOR3 activity or levels.
 XX
 XX Sequence 403 AA;
 CC
 CC Query Match 31.5%; Score 733; DB 20; Length 403;
 CC Best Local Similarity 39.4%; Pred. No. 9.8e-70;
 CC Matches 164; Conservative 60; Mismatches 114; Indels 78; Gaps 9;
 QY 1 MATHKALVMCLGLPLFLFPG-ANAQGHVPGCCSGLNPLYYNLCDSRGANGIVLEAVAGA 59
 DB 9 mrahgvltfil---lflvtsvasenastgrcgldilpqvsvicldaiwgivveavaga 65

